Department for Education

Research Report DFE-RR220

The impact of Sure Start Local Programmes on seven year olds and their families

The National Evaluation of Sure Start (NESS) Team

Institute for the Study of Children, Families and Social Issues, Birkbeck, University of London

This research report was commissioned before the new UK Government took office on 11 May 2010. As a result the content may not reflect current Government policy and may make reference to the Department for Children, Schools and Families (DCSF) which has now been replaced by the Department for Education (DfE).

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

The National Evaluation of Sure Start Team is based at the Institute for the Study of Children, Families & Social Issues, Birkbeck, University of London, Malet Street, London, WC1E 7HX

Core Team

Professor Edward Melhuish, Institute for the Study of Children, Families & Social Issues, Birkbeck (Executive Director)

Professor Jay Belsky, Human & Community Development, University of California, Davis: also Birkbeck (Research Director)

Professor Alastair Leyland, MRC Social & Public Health Sciences Unit, Glasgow (Statistician)

Impact Module

Professor Edward Melhuish, Institute for the Study of Children, Families & Social Issues, Birkbeck (Director)

Professor Angela Anning, Institute for the Study of Children, Families & Social Issues, Birkbeck (Investigator)

Professor Sir David Hall, University of Sheffield (Investigator)

Implementation Module

Professor Jane Tunstill, (Director)

Mog Ball (Investigator)

Pamela Meadows, National Institute of Economic & Social Research (Investigator)

Cost Effectiveness Module

Pamela Meadows, National Institute of Economic & Social Research (Director)

Local Context Analysis Module

Professor Jacqueline Barnes, Institute for the Study of Children, Families & Social Issues, Birkbeck (Director)

Dr Martin Frost, Birkbeck (Investigator)

Support to Local Programmes on Local Evaluation Module

Professor Jacqueline Barnes, Institute for the Study of Children, Families & Social Issues, Birkbeck (Director)

Data Analysis Team

Mark Hibbett, Institute for the Study of Children, Families & Social Issues, Birkbeck

Dr Andrew Cullis, Institute for the Study of Children, Families & Social Issues, Birkbeck

Acknowledgments

This report concludes the Impact module of NESS, which has been investigating the impact of Sure Start Local Programmes (SSLPs) upon young children and their families.

The research team wish to thank, in particular, the families and their children who have been wonderfully cooperative throughout this project. In addition the research team wish to thank the SSLPs, Local Authorities, Department for Education and other officials who have helped them in their work.

During the course of NESS the research team has been helped by an Advisory Group, whose input has been invaluable, and the research team thanks them for their input. The Advisory Group are:
Dr. Mitch Blair
Dr. Carolyn Davies
Professor Paul Gregg
Professor Sir Michael Rutter
Professor Marjorie Smith
Professor Sarah Stewart-Brown
Professor Kathy Sylva
Professor Peter Tyler
Professor Jane Waldfogel

Since 2001 NESS has produced 57 reports (including this report) for government and accompanying summaries. These reports and summaries can all be accessed at the website for NESS, www.ness.bbk.ac.uk. These reports are listed in Appendix H. In addition the research team have published their findings in peer-reviewed journals and book chapters, and a list of these publications is also available from the NESS website.

The work on the evaluation of the successors to SSLPs (Sure Start Children's Centres) is being carried on by the ongoing research project the Evaluation of Children's Centres in England, (see http://www.education.ox.ac.uk/research/fell/research/evaluation-of-children-centres-in-england-ecce/ and http://www.natcen.ac.uk/study/evaluation-of-children's-centres-in-england)

1. HM Treasury. Modern public services for Britain investing in reform. Comprehensive spending review: new public spending plans. London: HMSO, 1998.

THE IMPACT OF SURE START LOCAL PROGRAMMES ON SEVEN YEAR OLDS AND THEIR FAMILIES

Report of the Longitudinal Study of 7-year-old Children and Their Families

EXECUTIVE SUMMARY

Background

The ultimate goal of Sure Start Local Programmes (SSLPs) was to enhance the life chances for young children growing up in disadvantaged neighbourhoods. Children in this type of neighbourhood are at risk of doing poorly at school, having trouble with peers and agents of authority (i.e., parents, teachers), and ultimately experiencing compromised life chances (e.g., early school leaving, unemployment, limited longevity). The children in SSLP areas have low cognitive and language development being on average roughly one standard deviation below the population mean (NESS 2005). This indicates that the average child in an SSLP area is functioning at about the level of the 15th percentile. This represents a very substantial developmental handicap. These factors have profound consequences not just for the children but for their families, communities, and for society at large. Thus, SSLPs not only aimed to enhance health and well-being during the early years, but also to increase the chances that children would enter school ready to learn, be academically successful in school, socially successful in their communities and occupationally successful when adult. Indeed, by improving the developmental trajectories of young children at risk of compromised development, SSLPs aimed to break the intergenerational transmission of poverty, school failure and social exclusion. Such a strategy was a profound innovation for policy in the UK.

SSLPs were strategically situated in areas of high deprivation and they represented an innovative intervention unlike almost any other aiming to enhance the life prospects of young children in disadvantaged families and communities. One distinguishing characteristic was that the programme was *area based*, with *all* children under five

years of age and their families living in a prescribed area serving as the "targets" of intervention. This was seen as having the advantage that services (e.g. childcare, family support) within a SSLP area would be universally available, thereby avoiding stigma that may accrue from individuals being targeted. In the early years of SSLPs, by virtue of their local autonomy and in contrast to more narrowly-defined early interventions, SSLPs did not have a prescribed "curriculum" or set of services, especially not ones delineated in a "manualised" form to promote fidelity of treatment to a prescribed model. Instead, each SSLP had extensive local autonomy over how it fulfilled its mission through service delivery.

From 2005 to 2006, fundamental changes were made in SSLPs, as they came under the control of Local Authorities and were operated as Sure Start Children's Centres. This modified service delivery by making the guidelines for children's centres more specific about the services to be offered. Nonetheless there was still substantial variation among Local Authorities and areas within Local Authorities in the way the new children's centre model was implemented. This continued to pose challenges for evaluating their impact, as each SSLP or children's centre remained unique.

Evaluating SSLP Impact

In assessing the impact of SSLPs on child and family functioning over time, the impact study of the National Evaluation of Sure Start (NESS) has followed up over 5000 7-year-olds and their families in 150 SSLP areas who were initially studied when the children were 9 months, 3 and 5 years old. The 7-year-old study followed up a randomly selected subset of the children and families previously studied at younger ages.

The comparison group of non-SSLP children and their families, against which the NESS sample was compared, was selected from the entire Millennium Cohort Study (MCS) cohort. Their selection was based upon identifying and selecting children living in areas with similar characteristics to SSLP areas, but which did not offer SSLP services. This enabled comparisons amongst children and families from similar areas in order to detect possible effects of SSLPs on children and families.

Methodological Issues

Every study will have some methodological constraints. Here we summarise those that apply to this study and how we tried to mitigate them.

The data for the NESS and comparison samples of 7-year-olds and their families were collected two years apart and by two different research teams. This makes attributing SSLP effects to SSLP exposure per se difficult as they could potentially reflect changes in communities or society more generally across the two-year period or be the result of differences in measurement by the two research teams. We tried to mitigate this latter effect through close cooperation between teams and cooperation in staff training.

However, it should be noted that measures of educational achievement deriving from Key Stage 1 (KS1) assessments are free from problems linked to time of measurement or the differences between research teams in that KS1 assessments were completed by teachers independent of any research team, according to national measurement guidelines. Also the standardisation (or equivalisation) of KS1 scores by year of measurement further ensures the comparability of data across studies and years of measurement.

Missing data are unavoidable in a longitudinal study of this size, i.e., data that were not collected either because families could not be contacted or because of the decision of the funders not to follow up all those seen at 5 years of age when they were age 7. In order to counter possible bias due to missing data, comparisons between the 7-year-olds and their families participating in the SSLP and the comparison group were conducted for three different but overlapping samples:

1. Those children/families interviewed at age 7 for both SSLP and the non-SSLP sample for whom *complete data* were available (i.e. no missing data whatsoever on measurements used in this report). These cases numbered 3,282 in the SSLP sample and 1,127 in the non-SSLP sample, but eliminating cases with missing data may result in non-random loss of data and possibly biased results.

To compensate for this possibility two samples for analysis used imputation to replace missing data:

2. All those seen at age 7 whether or not there was complete data at age 7 (N=3,558 for SSLP, 1,436 for non-SSLP).

3. Those seen at 3 years old regardless of whether they were also seen at 7 years old (N=5,883 for SSLP, 1,879 for non-SSLP).

Imputation allows estimation of data for those lacking measurements on a variable by using all the other available information on all individuals. In essence, it uses the statistical relations among all variables to calculate what a missing value might be, while taking into consideration the likelihood of error in such estimates.

Given that results could differ across these analyses and that each approach has both strengths and weaknesses, the decision was made before any analyses were conducted that only SSLP effects (i.e., SSLP vs. non-SSLP) that proved significant across *all three sets* of analyses would be regarded as reliable and thus meaningful for presentation and interpretation in this report. This approach was taken as results that were significant across all three sets of analyses are the most reliable and suitable for policy planning.

Key Findings

After taking into consideration pre-existing family and area background characteristics, (as well as primary school characteristics for child outcomes), the three sets of analyses comparing children and families living in SSLP areas and those living in similar non-SSLP areas revealed some beneficial SSLP effects related to family functioning and maternal well-being. This was the case when effects were evaluated with respect to child/family functioning when the children were age 7 and with respect to change over time in child/family functioning from age 3 (or 9 months for worklessness) until age 7.

The Impacts of SSLPs When the Children Were Aged 7

Evaluation of SSLP effects involved 15 child/family functioning outcomes at age 7. As well as considering across the board effects, we also considered whether effects applied to children/families within SSLP sub-populations.

After taking into consideration pre-existing family, area and school characteristics, four positive effects of SSLPs emerged from these 15 outcomes at age 7, two of which applied to the whole population and two of which applied to sub-populations. For the whole population, mothers in SSLP areas compared to their counterparts not living in SSLP areas, reported:

(1) engaging in less harsh discipline;
(2) providing a more stimulating home learning environment for their children;

additionally for sub-populations, mothers in SSLP areas reported:

(3) providing a less chaotic home environment for boys (not significant for girls);
(4) having better life satisfaction (lone parent and workless households only).

The results for sub-populations can be as important as those for the total population, and knowing about sub-population differences can inform the targeting of services. This is increasingly important as children's centre services are often targeted on the most vulnerable, and also service delivery may be targeted differently for specific sub-populations.

There were no consistent differences on any of the four child educational development outcomes, or on four child social and behavioural outcomes or on the two child health outcomes.

In summary, significant positive effects of SSLPs emerged for four out of 15 outcomes, two of which applied to the population overall and two to sub-populations. These effects, for family functioning and maternal well-being[1], appeared to apply across the full range of SSLPs regardless of level of deprivation.

[1] Definitions of the outcomes can be found in Appendix C.

What is the Effect of SSLPs on Progress over Time?

Additional evidence of positive SSLP effects for the whole population emerged on three of eight repeatedly-measured outcomes when the focus of evaluation was on change in parent and child functioning between 3 and 7 years (or between 9 months and 7 years for workless household status). Mothers in SSLP areas relative to those residing in comparison areas:

(1) showed a greater improvement in the home learning environment;
(2) reported a greater decrease in harsh discipline;

additionally for sub-populations, mothers in SSLP areas reported:

(3) greater improvement in life satisfaction (lone parent and workless households only) than counterparts not in SSLP areas.

Note that there were only beneficial effects and no negative effects discerned in any of the analyses.

Are These Effects Applicable Across All 150 SSLPs?

It is important to know whether effects apply across all SSLP areas regardless of deprivation. This was checked by analysing differences in effects across all SSLPs and such analyses indicated that effects were similar regardless of level of area deprivation. Since the Millennium Cohort Study (MCS) does not include many children from the most extremely disadvantaged areas (stratum 5), NESS focused the comparative analysis on the other disadvantaged areas that are still within the 20% most disadvantaged areas in the country. In order to consider whether there might be a different impact for children in stratum 5, the outcome data of the Sure Start children in the other disadvantaged areas (strata 2-4) were compared with the outcome data for the Sure Start children in the most extremely disadvantaged areas (stratum 5), taking account of the differences in other child, family, area and school variables that might affect outcomes. Children and families in stratum 5 appeared to be developing equivalently to those in strata 2-4 when taking their extra disadvantage into account. This suggests that the most disadvantaged groups benefitted in an equivalent way to the other groups.
This finding is consistent with the interpretation that SSLP effects should generalise across all 150 SSLPs in this study, all of which were drawn from the first four waves of SSLPS, which were originally funded from before 2002.

Pre-school and Primary School Education

The main evidence for population-wide early years programmes affecting child development stems from research on the effects of high quality pre-school education, which has been found, repeatedly, to be associated with improved cognitive and social development (e.g. Sylva et al., 2010; Vandell et al., 2010; Melhuish 2011). While pre-school education was (and remains) a major part of what SSLPs (now children's centres) offered, it would also have been available to children in non-SSLP areas. Since 2004 free part-time pre-school provision has been available to every child from three years of age, and 95% of eligible children take up this offer (Statistical First

Release, DfE June 2011[2]). Hence there are unlikely to be differences in the pre-school education experienced by the SSLP and non-SSLP samples, particularly as the quality of pre-school provision is equivalent in SSLP and non-SSLP areas (Melhuish et al., 2010a). This equivalence of amount and quality of pre-school education experience across those living in SSLP and non-SSLP areas could well be responsible for the failure to detect SSLP effects on children at age 5 (apart from physical health measures) and at age 7. That is, it *could* be that developmental advantages that were identified for children in SSLPs at age 3 were not detected at age 5 (NESS, 2010) or at age 7 because almost all children, whether in SSLP areas or not, had access to pre-school education, which resulted in "catch up" for those children in non-SSLP areas.

In addition at the time of the 7-year-old assessments all children would have had approximately 3 years of primary school experience, which may also have contributed to an equalising of overall experience likely to affect development between children in the two study groups, given that they have had equivalent pre-school education experience. This interpretation is consistent with the evidence that high quality early childhood education and care (ECEC) (e.g. Melhuish, 2011) will shift the population curve for child outcomes, and this is the only type of early intervention for which evidence is currently available for shifting the population curve through enhancing the development of all children in the relevant population, rather than lifting the "tail" of the population through targeted intervention. Other types of intervention strategy have been found to be successful in helping targeted small groups of families with young children, notably well-implemented "manualised" parent support programmes (Barrett, 2010). Such programmes are sometimes used by Sure Start programmes, and they can help lift the tail of the population. Also there are examples of childcare in the first 3 years combined with parent support (e.g., Early Head Start; Love et al., 2002) that have also been found to have substantial impact for disadvantaged populations.

Conclusions

The NESS research team faced methodological challenges and these are outlined here and in more detail later in the main report. These issues place some limits on the study's ability to afford strong causal inferences about effects of SSLPs on children and families. Early decisions not to undertake a randomised control trial and to double the number of SSLPs (reducing the opportunity to identify suitable comparison areas) meant that the evaluation had to use the MCS cohort as a source of comparison data. This resulted in a two year gap between SSLP and comparison data such that any SSLP-comparison group differences might be due to time effects. This limitation was overcome by the research team with respect to education outcomes such as KS1 scores. Despite methodological challenges a great deal has been learnt over the years from the many reports provided by NESS, and these have clearly influenced policy. This report shows how some of those changes have had a positive impact, for instance on improving the reach and impact on the most vulnerable families, and there are no longer any negative effects associated with Sure Start programmes. While bearing in

[2] DfE, Statistical First Release 22nd June 2011: http://www.education.gov.uk/rsgateway/DB/SFR/s001009/index.shtml

mind the methodological caveats, it is possible to draw the following conclusions from this latest phase of the impact study, which shows only beneficial effects.

In summary, significant effects of SSLPs emerged for four out of 15 outcomes, two of which applied across the board and two of which applied to sub-populations. These four outcomes were concerned with family functioning and maternal well-being. For the whole population, mothers in SSLP areas relative to their counterparts in non SSLP areas reported:

(1) engaging in less harsh discipline;
(2) providing a more stimulating home learning environment for their children;

additionally for sub-populations, mothers in SSLP areas reported:

(3) providing a less chaotic home environment for boys (not significant for girls);
(4) having better life satisfaction (lone parent and workless households only).

Additional evidence of positive SSLP effects emerged for three of eight repeatedly-measured outcomes when the focus of evaluation was on change in parent and child functioning between 3 and 7 years (9 months and 7 years for workless household status). Mothers in SSLP areas relative to those residing in comparison areas:

(1) showed a greater improvement in the home learning environment;
(2) reported a greater decrease in harsh discipline;

additionally for sub-populations, mothers in SSLP areas reported:

(3) greater improvement in life satisfaction (lone parent and workless households only) than counterparts not in SSLP areas.

The effects for lone parent and workless households can be regarded as evidence that SSLPs were being successful in affecting 'hard to reach' groups within SSLP areas, which in the early stages of SSLPs had appeared to be challenging.

No consistent SSLP effects for child development emerged at 7 years. This is likely to be due to high levels of participation in the 3 and 4 Year Old Free Entitlement to pre-school education across England, which has resulted in most of the MCS children also benefitting from early years learning opportunities. Therefore the pre-school education experiences shown to be important for educational outcomes in other research (Sylva et al. 2010; Melhuish et al., 2008b) are much the same for both SSLP and non-SSLP groups. Additionally, by age 7 children were in their third year of primary school and so they have 3 years of primary school in common as well, which may partly account for the similarity in child outcomes across the study groups. It might also be related to the fact that, as reported by the Implementation module (Tunstill et al., 2005), in the earlier phases of Sure Start, services in many SSLPs focussed more on providing support and intervention for parents, than on programmes that target child daily experiences most likely to boost child development.

In summary, SSLPs have beneficial effects as follows:

i) on family functioning and maternal well-being that have persisted until the children were age 7, and some improvements have continued over time for some measures;

ii) but there was no impact on child outcomes, which is likely to be, at least in part, because of the introduction of universal free early education for all children whether in Sure Start areas or not.
It is noteworthy that (1) the discerned SSLP effects appear to apply to all areas served by SSLPs regardless of level of deprivation, and (2) apply to all children and families regardless of family deprivation. This indicates that since the earliest stages of the evaluation SSLPs have moved forward and now appear to be engaging the most vulnerable groups in the most deprived areas. This suggests that SSLPs really did respond to the early findings, and remedy the problems with engaging vulnerable groups.

The results discerned in this follow-up study of 7-year-old children and their families provide some support for the view that government efforts to support children/families via the original area-based approach to Sure Start paid off to some degree with parent outcomes, but not with regard to child outcomes. Since its early days Sure Start has evolved considerably, responding to research findings and both internal and external feedback. In particular, policy developments have clarified guidelines and worked to strengthen service delivery. However, at the same time, one cannot entirely discount the possibility that apparent effects are an artefact of the two-year gap between SSLP and non-SSLP data collections. It is possible that the results might have been more positive or more negative if this two-year gap in data collection did not exist. Nevertheless, while the results are modest they do indicate value in the work of Sure Start Children's Centres. Importantly, children's centres have been found to be immensely popular with parents and, as demonstrated by this evidence, they have been successful in reaching the parents who are likely to be the most disadvantaged. The success of SSLPs in engaging and supporting the poorest families without stigma means they provide an infrastructure that is well placed to engage the most vulnerable groups and support them effectively. Also the beneficial effects for parents are persisting at least two years after their last contact with Sure Start programmes; often social interventions do not have such a sustained impact and can suffer from "wash-out".

It is possible the beneficial effects on parenting and more effective home learning environments may produce improved child outcomes when the children are older. Other interventions that have affected parenting report long-term beneficial effects upon adolescent criminality and substance abuse (Olds et al., 1999). It may be that the parenting effects discerned by age 3 are too little too late to produce the improvements necessary for language development that could affect subsequent educational and social improvement. If child outcomes are to be enhanced the established Sure Start infrastructure needs to give greater emphasis to services that will improve child outcomes, particularly language development and children's daily experiences, for the children served. In relation to this last point another NESS report (Melhuish et al., 2010a) has shown that when Sure Start programmes provide high quality pre-school childcare, child language development can improve, and other research shows that improvements can be long-lasting and extend to educational and social outcomes. All this suggests that Sure Start Children's Centres are well-placed to provide improved

integrated services that will help support the most disadvantaged children and families and in a way that can contribute to narrowing the gap between the children of disadvantaged and more advantaged families. However, they will need to focus more directly on improvements to young children's daily experience, which is a primary engine of child development, if they are to improve child outcomes.

CONTENTS

APPENDICES

TABLES

FIGURES

1. INTRODUCTION

In 1998 the Cross-Departmental Review of Services for Young Children concluded that disadvantage among young children was increasing and that early intervention could prevent poor outcomes (HM Treasury, 1998a). It also noted that current services were uncoordinated and patchy and recommended there be a change in service design and delivery, suggesting that programmes should be jointly planned by all relevant bodies, and be *area-based*, with *all* children under five and their families in an area being clients. In July 1998, the then Chancellor of the Exchequer, Gordon Brown, introduced Sure Start aimed at providing quality services for children under five years old and their parents (HM Treasury, 1998b). The original intent of the programme design was to focus on the 20% most deprived areas, which included around 51% of children in families with incomes 60% or less than the national median, i.e. the official poverty line (Melhuish & Hall, 2007).

The ultimate goal of Sure Start Local Programmes (SSLPs) was to enhance the life chances for young children growing up in disadvantaged neighbourhoods. Children in these communities were at risk of doing poorly at school, having trouble with peers and agents of authority (i.e., parents, teachers), and ultimately experiencing compromised life chances (e.g., early school leaving, unemployment, limited longevity). This has profound consequences not just for the children, but for their families, communities, and for society at large. Thus, SSLPs not only aimed to enhance health and well-being during the early years, but to increase the chances that children would enter school ready to learn, be academically successful in school, socially successful in their communities and occupationally successful when adult. Indeed, by improving, early in life, the developmental trajectories of children known to be at-risk of compromised development, SSLPs aimed to break the intergenerational transmission of poverty, school failure and social exclusion.

SSLPs were unlike almost any other intervention devoted to enhancing the lives of young children growing up in disadvantaged families and communities. What made the initiative so different was that it was *area based*, with *all* young children and their families living in a prescribed area serving as the “targets” of intervention. In contrast to more targeted interventions carried out in the USA, SSLPs initially did not have a prescribed “curriculum” or set of services. Thus the programmes were not “manualised” in a form that would promote fidelity of treatment to a prescribed model. Instead, each local programme was charged with improving existing services and creating new ones as needed, without specification of how services should be changed. This was in contrast to early interventions previously demonstrated to be effective, be they childcare based, like the Abecedarian Project (Ramey et al., 2000); home based, like the Nurse Family Partnership, (Olds et al., 1999); or even a combination of centre and home based, like Early Head Start (Love et al., 2002).

From 2005-2006 onwards SSLPs came under Local Authority control and were expected to implement a children’s centre model. As the administration of Sure Start

Children's Centres changed and the guidelines were more specific about the services to be offered, SSLPs changed the nature of their services. Nonetheless, the guidelines were still not so specific as to homogenise the services being delivered or how well they were delivered. There remains substantial variation across Local Authorities and between areas within Local Authorities in the way the children's centre model is implemented. Thus in contrast to other, more highly specified, early interventions, SSLPs and subsequently children's centres are much more varied in terms of what they deliver and how they deliver it. This has posed challenges to evaluating their impact, as each programme is relatively unique.

Further changes to children's centres have been occurring as a result of the change in government in May 2010, and the cuts in public spending that started to have an impact from April 2011. The consequences of these changes are yet to be fully realised. However it is still the core purpose of Sure Start Children's Centres to improve outcomes for young children and their families, with a particular focus on the most disadvantaged, so children are equipped for life and ready for school, no matter what their background or family circumstances. Children's centres are still seen as key to fulfilling this purpose but LAs are looking to achieve greater efficiencies by merging children's centres and clustering others under single management teams. The ongoing Evaluation of Children's Centres in England[3] study may throw light on these changes.

Given their ambitious goals, it is clear that the ultimate effectiveness of SSLPs cannot be determined for some time and that children growing up in communities with SSLPs would need to be studied well beyond their early years before a final account of the impact of SSLPs will prove possible. Nevertheless, by studying children and families in SSLPs during their early years, it may well prove possible to detect evidence of early effectiveness. The longitudinal phase of the impact study of the National Evaluation of Sure Start (NESS) has built upon the first, cross-sectional phase (NESS, 2005) and was designed with this goal in mind. Specifically, over 5000 children growing up in 150 SSLP areas and first studied, along with their families, at 9-months, 3 years and 5 years of age have been studied again when 7-years-old. In order to evaluate the effects of SSLPs on child and family functioning, the SSLP children/families are compared with similar children/families participating in the Millennium Cohort Study (MCS) who have also been studied at 9 months, 3, 5 and 7 years of age. Selection of comparison children/families from the MCS was based upon their residing in similar areas to those of the NESS longitudinal sample, but not benefiting from an SSLP.

Early cross-sectional comparisons of 9-month-olds and 3-year-olds and their families in 150 SSLP areas with counterparts living in 50 communities destined to become SSLP areas, revealed some small effects of SSLPs on child/family functioning (NESS, 2005;

[3] http://www.education.ox.ac.uk/research/fell/research/evaluation-of-children-centres-in-england-ecce/ and http://www.natcen.ac.uk/study/evaluation-of-children's-centres-in-england

Belsky, Melhuish, Barnes, Leyland, Romaniuk, & the NESS Research Team, 2006). Differences between these two sets of families indicated, principally among the 3-year-olds and their families that most of the children and families living in SSLP areas benefited somewhat from the programme.

Overall, benefits were that 9-month-olds experienced less household chaos and mothers of 3-year-olds proved more accepting of their children's behaviour (i.e. less slapping, scolding, physical restraint). Mothers of 3-year-olds who became parents after their 20th birthday (86% of sample) engaged in less negative parenting when living in SSLP areas rather than the comparison communities. Also their 3-year-old children exhibited fewer behaviour problems and greater social competence when living in SSLP communities than in comparison communities, and evidence indicated that these effects for children were mediated by SSLP effects on the parenting of non-teen mothers (i.e., greater acceptance, less negative parenting). However the most disadvantaged children/families (i.e., teenage mothers, workless or lone parent households) seemed to experience some adverse effects of living in SSLP areas Adverse effects for children of teen mothers (14% of sample) were that they scored lower on verbal ability and social competence and higher on behaviour problems than their counterparts in comparison areas. Children from workless households (39% of sample) and children from lone-parent families (36% of sample) also showed evidence of adverse effects of SSLPs, scoring significantly lower on verbal ability when growing up in SSLP areas than did their counterparts in comparison communities.

A follow-up study at 3 years of age of the 9-month-olds from the initial cross-sectional study presented a substantially different picture of the effects of SSLPs, in showing only beneficial impact (NESS, 2008, Melhuish et al., 2008a). These children, at 3 years of age, were compared with similar children/families participating in the Millennium Cohort Study (MCS), living in similar areas not receiving SSLPs, and who were also studied at 9 months and 3 years of age. After allowing for pre-existing family and area background characteristics, a variety of beneficial effects associated with living in SSLP areas emerged (on 7 of 14 outcomes assessed) and there was no evidence that the beneficial effects varied by subgroup or that adverse programme effects were identifiable for any sub-populations. More specifically, children growing up in SSLP areas showed better social development, exhibited more positive social behaviour and greater independence/self-regulation than their non-SSLP counterparts.

These beneficial SSLP effects may well have been the result of the better parenting that was also associated with living in SSLP areas. Parents in SSLP areas showed less negative parenting while providing their children with a better home learning environment than parents residing in non-SSLP areas. Finally, these beneficial effects of SSLPs on children and families may themselves have been a function of the greater use of support services reported by parents living in SSLP areas relative to those not living in such areas. Parents in SSLP areas reported using more services than the comparison group of parents. In addition, children in SSLP areas were more likely to have received the recommended immunisations and were less likely to have had an accident-based injury in the year preceding assessment. These latter two results

(immunisations and accidents) may have been an artefact of time-of-measurement effects, however, in that the MCS sample was born, on average, two years before the NESS sample and the two outcomes in question showed evidence of more favourable scores the later in time that data collection took place in both of the samples. This confounding of time with the two outcomes in question raised the possibility that time-of-measurement rather than growing up in SSLP areas accounted for these (apparent) SSLP effects.

The fact that detected effects of SSLPs in the second phase of the NESS impact study did not vary by population subgroups was markedly different from the first phase of evaluation. Earlier evidence was that the most disadvantaged 3-year-old children and their families (i.e., teen parents, lone parents, workless households) were doing less well in SSLP areas, while somewhat less disadvantaged children and families benefited (i.e., non-teen parents, dual parent families, working households). The subsequent evidence collected when children were age 3 years revealed benefits for all sections of the population served.

Various explanations could be offered for the differences between the 2005 and 2008 findings. Although it was not possible to entirely eliminate methodological explanations, it seemed reasonable that the contrasting results accurately reflected the contrasting experiences of children and families in SSLP areas in the two phases. Whereas the 3-year-olds in the first phase were exposed to 'immature' programmes—and probably not for their entire lives—the 3-year-olds and their families in the second phase were exposed to better developed programmes throughout the entire lives of the children. Also programmes had the opportunity to learn from the earlier phase of the evaluation, especially with respect to making greater efforts to reach the most vulnerable households. Thus differences in the amount of exposure to programmes and the quality of SSLPs may well have accounted for both the initial adverse effects detected for the most disadvantaged children and families and the subsequent beneficial effects discerned for almost all children and families living in SSLP areas.

When the children in the longitudinal sample in SSLP and non-SSLP areas were followed up at 5 years of age, the results again were largely indicative of beneficial effects associated with living in SSLP areas (NESS, 2010). Across 21 outcomes, 8 revealed significant effects of SSLPs, 6 of which were positive and 2 negative. The main impacts identified for children were that those growing up in SSLP areas had lower BMIs than children in non-SSLP areas. This was due to their being less likely to be overweight with no difference for obesity. They also experienced better physical health than children in non-SSLP areas. No differences emerged between the SSLP and non-SSLP groups on 7 measures of cognitive and social development from the Foundation Stage Profile completed by teachers, 4 measures of socio-emotional development based on mothers' ratings, and mothers' ratings of area safety.

The positive effects for maternal well-being and family functioning, in comparison with those in non-SSLP areas were that mothers in SSLP areas reported providing a more cognitively stimulating home learning environment for their children, a less chaotic home

environment and engaged in less harsh discipline. They also reported greater life satisfaction. On the negative side, however, in comparison with those in non-SSLP areas, mothers in SSLP areas reported more depressive symptoms and parents in SSLP areas were less likely to visit their child's school for parent/teacher meetings or other arranged visits. It should be noted, however, that the overall incidence of such visits was low generally for all study participants.

In the current report children and families who were previously seen at 9 months, 3 and 5 years of age in the NESS (SSLP) or MCS (non-SSLP) longitudinal studies are compared to determine whether differences in child and family functioning found at 3 and 5 years of age persist until 7 years of age, and whether any other differences emerge. Efforts have been made, when equivalent measurements were taken at 3- and 7-years of age, to see if SSLP and non-SSLP children differed in terms of the developmental change they manifest across this four year period. At this fourth phase of the NESS impact study the children are in their third year of primary school. Hence it was important to include some measures of the primary schools that the children attend. Thus data used in our analyses derive from child assessments, parental interview, Ofsted and DfE data on primary schools, and from National Assessments at the end of Key Stage 1.

2. RESEARCH DESIGN

2.1 Background

When, in 2000, the government decided to almost double the number of SSLPs from 260 to more than 500, the decision was made to rely upon the MCS to provide a comparison sample, as finding another comparison group would have been impractical. For this reason, the NESS impact study has sought to ensure that its procedures, methods and measurements mirrored, for the most part, those in the nationally representative Millennium Cohort Study (MCS).

Several alternative strategies for using the MCS sample and data were initially considered. One, for example, was to rely upon all the children/families participating in the MCS and statistically control for any differences within and across samples on a host of child, family and community background factors. A second called for using as a comparison only disadvantaged children/families living in areas of concentrated deprivation, thereby maximising family and community similarity to SSLP families and communities.

Since the start of the NESS impact study in 2001 several methodological advances have occurred in the study of environmental influences on child and family functioning, though they have a longer history in other fields of inquiry. Many of these advances involve statistical procedures and ways of accounting for potential pre-existing differences between groups that vary on an independent variable of interest, like SSLP exposure, especially with respect to omitted variables, that is, variables that might be important yet have gone unmeasured (McCartney, Bub & Burchinal, 2006). One of these advances is “propensity scoring”, which has been adopted in this study. Propensity Score Analysis (Rosenbaum & Rubin, 1983; Rubin, 1997; Pearl, 2009) can be used to address selection bias, in this case the possibility that those who experience a treatment (i.e., SSLP) may differ in unmeasured ways from those who do not, when randomisation is not possible. The term propensity refers to “a conditional probability of an individual being in a treatment group, given a set of background variables for that individual” (McCartney et al., 2006, p. 114). In this study whether a child is in the treatment group is determined by whether or not the child lives in a SSLP area; the problem therefore reduces to identifying those areas that have a greater or lesser propensity of having populations that are similar to those of SSLP areas.

Propensity scoring estimates the likelihood of being a SSLP area by distinguishing between groups on area characteristics. The Local Context Analysis module of NESS developed a number of techniques that maximised the usefulness of data from diverse sources that predated the NESS impact study and that could be used for this purpose (see Barnes, 2007; Frost & Harper, 2007). We have used such techniques to provide detailed data on areas with the constraint that equivalent data must be available for MCS areas. Using such area-level data, 138 disadvantaged comparison areas were

initially identified with no geographic overlap with SSLP areas and which included MCS children. Of these, 72 MCS areas proved from a propensity score analysis to be suitable for comparisons with SSLP areas (NESS, 2008; Melhuish et al., 2008a). The 72 non-SSLP areas included 1,879 children participating in the MCS, who were seen at both 9 months and 3 years of age, and for whom there were adequate data for use in statistical analyses. Fuller details of the use of propensity scoring in selecting comparison areas are shown in Appendix A.

2.2 Methodological Issues

Some fundamental methodological issues arising from the nature of the comparison group and other elements of the study constrain the study's ability to address the core issue of effects of SSLPs on children/families.

1. Study design
2. Choice of a comparison group
3. Time of measurement of data
4. Parent report as a source of data
5. Cognitive and language development measures

1. Study design: Randomised Controlled Trials (RCTs) are often referred to as the 'gold standard' for evaluation methodologies. It is widely recognised that where RCTs are appropriate and well-executed they provide the strongest form of evidence, and allow the strongest inference with regard to causal attribution. Amongst their advantages RCTs solve the problem of selection bias through random assignment of the intervention. Those randomly selected for the intervention constitute the treatment (experimental) group and those not selected constitute the control group. After treatment has occurred differences in outcome between the treatment and control groups provide a measure of the effect of the treatment. An individually based RCT would not be appropriate for an intervention targeted at areas rather than individuals (such as SSLPs). On the other hand, a RCT based on randomisation of areas would have been possible, but the early roll-out and rapid expansion of the Sure Start programme precluded this as an option.

As a RCT was not possible, the NESS team selected the next best evaluation design based upon quasi-experimental methods. In this approach child and family outcomes are analysed as a function of whether participants are in a SSLP area or not, controlling for a range of covariates of child, family and community characteristics. This strategy provides an answer to the question of whether SSLPs have an effect after allowing for effects of child, family and community characteristics. Critics could argue that other unmeasured differences (e.g. genetic factors) may nevertheless affect the results. Even though the evaluation statistically controls for many relevant covariates, this criticism, which applies to all quasi-experimental research, can never be completely discounted.

2. The comparison group: the decision to use the Millennium Cohort Study (MCS) cohort as the source of a comparison, the non-SSLP group, had consequences for issues of time of measurement (see below), for variables that were chosen and how

they were measured, and also for which control and outcome variables could be used. Only variables measured in an equivalent manner in both MCS and NESS studies could be used in detecting SSLP effects. In order to facilitate the collection of equivalent data the NESS team has liaised throughout the project with the MCS team, though this has not guaranteed, to the extent originally desired, that measurement equivalence was maintained.

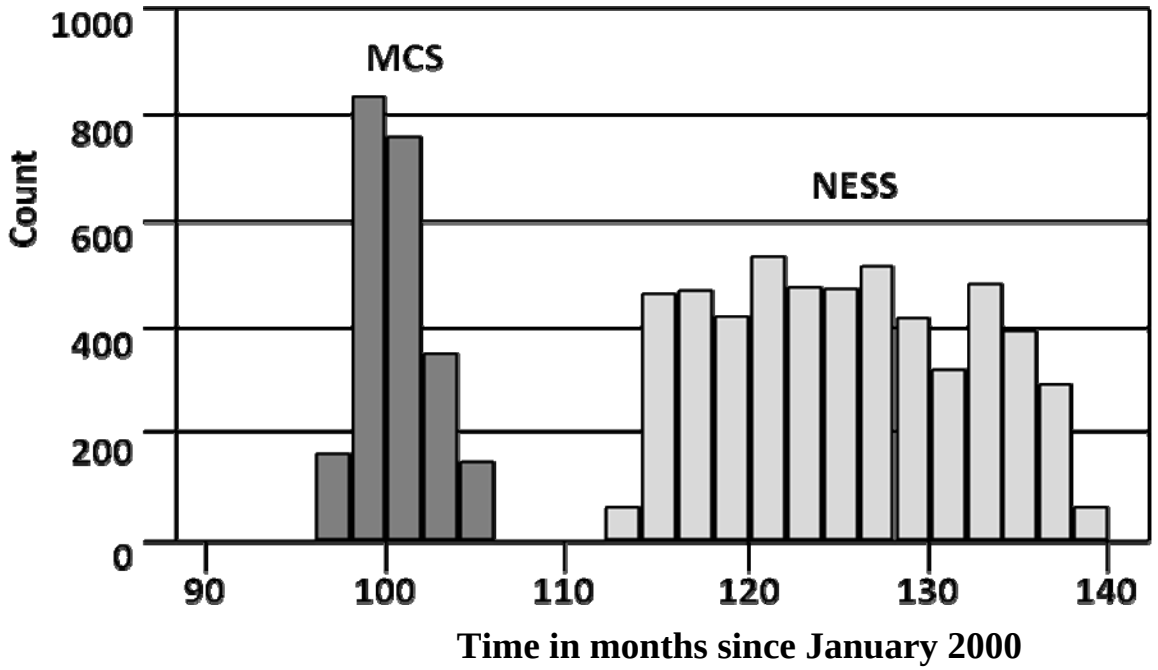
An additional complicating factor with respect to the MCS as the source of a comparison group is that it did not include many economically disadvantaged families residing in communities with the extent of concentrated disadvantage that characterised the SSLP areas. This meant that when it came to making comparisons involving MCS cases to evaluate SSLP effects, children/families in SSLP areas most characteristic of SSLP areas, with the highest levels of disadvantage, could not be included. Consequently the primary comparisons were less than ideal with respect to drawing conclusions about SSLP effects. How this situation was discovered and the manner with which it was handled is described in detail in section 2.5 below.

3. Time of measurement variations: Partly because of the time it took to get SSLPs "bedded down" and the desire to evaluate effects of "bedded down" SSLPs, the NESS and MCS longitudinal studies were not launched at the same time. MCS 7-year-old fieldwork took place between January 2008 and December 2008, and NESS 7-year-old fieldwork took place between June 2009 and June 2011. Hence, there exists, on average, a two year gap between the time of data collection for the MCS (non-SSLP) and NESS (SSLP) samples. A strategy adopted in an earlier phase of NESS to deal with this problem was to include time of actual data collection—operationalised as elapsed months since January 2000--as a covariate in analyses to discount any effects of time before testing for SSLP effects. For the current report this strategy was problematic because there was no overlap in when the MCS and NESS samples were seen at age 7 years (see Figure 2.1). This means that including both SSLP status (i.e., NESS vs. MCS) and time of measurement in the same statistical model would lead to major problems of collinearity. Hence it was decided not to include time of measurement in analyses. Therefore, time of measurement cannot be ruled out as an alternative explanation for almost any NESS/MCS differences or lack of differences, and thus SSLP effects or non-effects discerned.

However in the case of child outcome data deriving from National Assessments for Key Stage 1 (KS1) the time of measurement issue was not relevant because the NESS team secured total national KS1 data. This enabled the team to standardise KS1 measurements within each year of measurement before comparing SSLP and non-SSLP samples. Because equivalent national data do not exist for any other measurements used in this report, such standardisation was not possible for the other outcomes to be evaluated. It is important to note that measures of child development deriving from Key Stage 1 (KS1) data are free from any differences between research teams in that KS1measurement is done by teachers according to National Assessment guidelines. The standardisation (or equivalisation) of KS1 scores by year of measurement further ensures the comparability of data across studies and years of

measurement. Similar points will apply to future use of Key Stage assessments such as possible comparisons at 11 years (Key Stage 2).

Figure 2.1: Time of measurement for data for 7-year-olds



Thus overall there is no way to discount time-related alternative explanations for any 7-year-old SSLP effects discerned for outcomes other than KS1 scores. Such time-related alternative explanations could include any general trend (e.g., changes in the economy) or a specific event (e.g., a change of government) that might occur in the two-year gap between data measurements across the two samples.

4. *Parental report data:* With the exception of child weight and height, and child cognitive and educational functioning, all the child/family outcome data used to illuminate SSLP effects derive from parental reports. Although parent report measures as used in this study have been found to be useful and valid in many studies we need to consider the possibility that such parental reports may suffer from inaccuracy or bias related to individual or group characteristics or experience of parents. It is possible that such problems may influence the results, although there is no obvious reason for such problems to affect one of the samples in this study more than the other, except if SSLP parents' exposure to services gives them better knowledge of what the "best" answer might be.

5. *Child cognitive and language development*: When the NESS impact study was originally planned the intention was to investigate effects of SSLPs on children's formally tested cognitive and language development at ages 3, 5 and 7. At age 5 inspection of data from the two studies raised doubts about the equivalence of data

across samples, despite efforts to maintain equivalence across studies through training fieldworkers in similar ways. Similar concerns were also present with age 7 data. It was concluded that these standardised assessment data could not be confidently used in comparison between the NESS and MCS samples because of possible measurement artefacts. Hence at 7 years of age the Key Stage 1 assessments are relied upon as proxy measures of child cognitive performance as these measures were collected by teachers, independent of the research teams.

2.3 Intention to Treat Analysis

SSLPs were community-based; everybody in the community was potentially a beneficiary of the programme. As in the original cross-sectional impact study (NESS, 2005), an "intention to treat" design was adopted in the evaluation of the impact of SSLPs. Such an approach does not focus only on those children and families identified as having used SSLP services, but rather on all children and families living in SSLP areas; i.e., the total population for whom the intervention was intended. This focus is appropriate because SSLPs had as their targets all children under 5 years of age in their area and their families. Thus 9-month old children and their families in SSLP areas were randomly sampled and followed up at 3, 5 and 7 years of age, so that they could be compared with children and families similarly randomly sampled in the MCS, but not residing in SSLP areas.

2.4 Identifying Potential Matched Areas

The characteristics of an area were critical to it being allocated a SSLP. Hence it was decided that the MCS comparison children should live in areas that were as similar as possible to the SSLP areas. This required matching areas where MCS children live with the SSLP areas in the NESS longitudinal study as far as possible. A fundamental challenge was identifying small geographical areas that included reasonable numbers of children in the MCS that could serve as comparison areas. Geographical analysis was used (see Barnes, 2007; Frost & Harper, 2007) to identify deprived areas containing MCS children/families that were as similar as possible to SSLP areas. Geographic Information Systems (GIS) were used to select potential areas and to extract data on them. The main indicator used to identify and select areas at the first stage was the overall score of the Index of Multiple Deprivation (IMD) 2004 (ODPM, 2004). The specification of areas was complicated by the fact that the design of the MCS was based on sampling within 1998 electoral wards meaning that there was no direct comparability between the areas used in the MCS sampling and the areas for which IMD 2004 and Census information were available. To overcome this problem, areas containing MCS children were identified using individual postcodes following strict guidelines specified by the ESRC longitudinal studies committee to prevent disclosure of personal information.

Initial tests were made using the IMD 2004 data to select wards that contained MCS children not overlapping with SSLP areas. These showed that wards selected in this way were less deprived than the SSLPs. Although some contained MCS children living in relatively deprived localities, the overall IMD scores reflected the fact that wards

were large and contained both deprived and relatively non-deprived localities. It was necessary, therefore, to delineate potential comparator areas using the smaller, more focused, Super Output Areas (SOAs) so that relatively deprived localities could be defined more clearly. GIS were used to select SOAs within the same deprivation range as SSLP areas. By using an intersection method, any SOA that overlapped with an SSLP area was excluded. Any area selected had to contain more than 9 MCS children.

In order to enhance the comparability of SSLP and non-SSLP areas we created a measure of the levels of affluence of the areas surrounding the MCS and SSLP areas, to serve as an indication of the neighbouring influence on an area and the degree to which it was an isolated area of deprivation. A rule-of-thumb 750 metre buffer was created around each area to represent typical walking distance. Postcodes within each buffer and for the internal areas were extracted and linked to income data (mean household annual income). From this, the following measures were calculated: (1) The ratio of the internal and external buffer weighted means for comparison between the two; (2) percent of households in the buffer whose mean household income was greater than the national average, thereby providing an indication of how affluent the surrounding population was; and (3) a measure of household income variation in the buffer zones. With these and IMD data in hand, it proved possible to identify 138 potential comparison areas that included MCS children/families, but did not have an SSLP.

2.5 Propensity Scoring

The fundamental assumption underlying matching on the propensity score is that if the two groups (SSLP and non-SSLP) are balanced on all known area covariates, they are likely also to be matched on unknown and unmeasured covariates not included in the propensity analysis. Propensity scoring (Rosenbaum & Rubin, 1983; Rubin, 1997; Pearl, 2009) can be used to estimate the contextual similarity to residing in an SSLP area based on area (rather than individual) characteristics (Hill et al., 2005). We can then create “treatment” and “control” groups matched on their propensity to be an SSLP area. First, the probability of an area having an SSLP, its propensity score, was estimated. This involved logistic regression with the area’s status, SSLP vs. non-SSLP, serving as the outcome to be predicted and several indices of area deprivation and other socio-demographic area characteristics used as predictors of area status (see Appendix A). This propensity score was used as a one-number summary of all the predictor variables for each area. Any imbalance across groups with respect to the confounding area covariates was used as a diagnostic of the adequacy of the propensity model and led to the creation of a refined propensity score and better balance. If important variables distinguishing SSLP and non-SSLP areas have not been omitted, the comparison of outcomes between SSLP and non-SSLP groups should have minimal bias due to the non-random allocation of SSLPs to areas.

In order to implement propensity scoring analysis, it was essential to determine which of the 138 aforementioned MCS areas were sufficiently comparable to the SSLP areas to be useful in an analysis. Therefore, they were compared with the 150 SSLP areas on 85 indices of deprivation and other area characteristics obtained from administrative

sources (see Appendix A for more complete reporting of Propensity Scoring data, analysis and decision making).

SSLP area populations were, in general, more disadvantaged than the potential comparison areas drawn from the MCS. This posed problems in making comparisons between roughly equivalent SSLP and non-SSLP groups in order to evaluate putative SSLP effects. To deal with this, the SSLP and non-SSLP areas were each divided into five subgroups—or—"strata" reflecting the extent to which they were likely, on the basis of their deprivation and demographic characteristics, to be chosen as an SSLP area. On the basis of such "propensity scoring", areas in stratum 1 had the lowest propensity to be chosen as a SSLP area, because they had the least deprivation, and those in stratum 5 had the highest propensity to be chosen as a SSLP area, because they had the most deprivation. There proved to be only a single non-SSLP (MCS) area that qualified as having the highest propensity (stratum 5) to be chosen as a SSLP area; this was due to the paucity of very disadvantaged families and areas in the MCS data set. In the SSLP (NESS) sample, however, the reverse proved to be the case. Whereas 55 SSLP areas qualified for stratum 5 due to high levels of area and family deprivation, only two SSLP areas met criteria for having the lowest propensity to be chosen as a SSLP area (stratum 1) due to few SSLP areas being relatively less disadvantaged economically and demographically. The differential distributions of non-SSLP (MCS) and SSLP areas across more and less disadvantaged areas and thus strata, displayed in Table 2.1, posed analytic challenges (see below).

Table 2.1: Distribution of SSLP and non-SSLP Areas Using Propensity Scores to Stratify Areas

Propensity score	Sure Start	Non-SSLP	
Stratum	**N Areas**	**N Areas**	**N children**
1	2	53	1,041
2	15	40	970
3	33	22	818
4	45	10	565
5	55	1	21
Total	150	126	3,415

Children who were in either the NESS or MCS studies at 9 months and 3 years of age were potential candidates for the sample whose data were to be analysed. The end result of the initial propensity scoring analysis is that we succeeded in identifying in the MCS a sample of 3,415 children nested in 126 areas that could be potentially used as a non-SSLP comparison group. However, because there were so few MCS children/families in stratum 5, and so few NESS children/families in stratum 1 it was judged necessary to eliminate these strata when making SSLP vs. non-SSLP comparisons for purposes of detecting SSLP effects (i.e., comparisons were based on strata 2, 3 and 4). This meant that there were 72 non-SSLP areas with 2,353 children suitable for SSLP vs. non-SSLP comparisons (i.e., in strata 2-4). Of the non-SSLP children in these 72 areas there were 1,879 children and families who were seen at 3 years of age suitable for including in analyses and of these 1,655 were seen at 5 years of age, and of these 1,436 were seen at 7 years of age.

In order to best evaluate the effects of SSLPs, a two-stage analysis plan was implemented. First we restricted the main SSLP vs. non-SSLP comparisons to children/families in areas from Strata 2-4. Second, following these comparisons, we sought to determine whether the NESS children/families that were excluded from the SSLP vs. non-SSLP comparison functioned in ways similar to the other NESS/SSLP children/families. In all these comparisons the background characteristics of children, families and areas are controlled. If they scored similarly on outcome measures, this would indicate, though not prove, that any detected effects of SSLPs (in the first stage of inquiry) should generalise to all NESS children/families. Should NESS children/families in stratum 5, the biggest group (37%) of SSLP areas, prove to function better than those in other strata within the NESS sample, allowing for background characteristics, this would indicate, but again not prove, that any detected beneficial effects of SSLPs might under-estimate benefits of living in an SSLP area. In contrast, if NESS children/families in stratum 5 functioned more poorly than those in other strata within the NESS sample, allowing for background characteristics, this would indicate that any detected beneficial effects of SSLPs might reflect over-estimates of positive SSLP effects. Fuller details of the use of propensity scoring are given in Appendix A.

2.6 Sample

As already noted, the sample of the NESS longitudinal impact study is a sub-sample of those originally studied in the 9-month group of the cross-sectional impact study (NESS, 2005). Potential cross-sectional study participants living in 150 SSLP areas were identified with the assistance of the Child Benefit Office of (initially) the Department for Work and Pensions and (subsequently) HM Revenue and Customs. 14,900 potential participants were randomly selected from the Child Benefit Register and of these a total of 12,575 9-month olds and their families were enrolled in the study, representing a response rate of 84.4%. The aim was to have at least 8,000 children/families in the longitudinal study when the children were 3 years of age. Of those seen at 9 months of age, 11,118 children/families from the 150 SSLP areas were randomly selected to be approached by a NESS fieldworker in order to collect data when the child was 3 years of age. Of these families 9,192 (82.7%) participated in the 3-year-old data collection. Of those not participating 388 refused (3.5%), 1,484 (13.3%) proved not to be contactable, often because they had moved and were untraceable; and 54 (0.5%) were not seen for diverse 'other' reasons. Thus data collection was completed for 9,192 children and families when the children were 3 years of age. At 5 years of age 8,000 of the children and families seen at 3 years of age were randomly selected to be followed-up. Of those approached, data was successfully collected on 7,258 children and families, representing a response rate of 91.6%. These children and families constituted the NESS longitudinal sample at 5 years of age. Of the 7,258 NESS children and families seen at 5 years of age 6,000 were randomly selected for follow-up at age 7. Of those approached data was successfully collected on 5,447 children and families, representing a response rate of 90.8%.

The NESS children and families seen at 9 months but not seen at 7 years were compared with those seen on both occasions, separately for strata 1-5, on a range of demographic variables. Comparisons of those not seen at 7 years relative to those seen at both ages of measurements revealed that on six indicators families not re-studied were significantly less advantaged than those in strata 1-4, but significantly more advantaged than those in stratum 5 (i.e. ethnicity, English first language, poverty, occupational status, lone parent and workless household). There were only two variables (i.e. teen mother and parent education) where the families not seen at 7 years were more disadvantaged than stratum 5. (see Appendix B). This indicates that, despite attrition, the profile of people lost to the study falls within the range of people still in the study so, despite losing them, the results are still generalisable to the SSLP population as a whole. Implications of these differences are considered in the results section 3.4.

MCS children/families were identified and recruited through a similar strategy by the MCS research team. As described earlier, 1,436 MCS children had been seen at 9 months, 3, 5 and 7 years of age and were categorised in strata 2-4. This 7-year-old non-SSLP sample represented a response rate of 87% of those seen at 5 years. These children came from areas that were matched—more or less—by means of propensity scoring to SSLP areas. In the non-SSLP sample there were also children and families seen at 9 months but not at 7 years and they were compared on demographic characteristics to those seen on both occasions. The families not seen at 7 years were more likely to be more deprived on several variables (i.e. English first language, teen mother, poverty, occupational status, lone parent, workless household, and parent education) than the non-SSLP subsample seen at both ages (see Appendix B for full comparisons). As described the decision was taken to test for differences between the SSLP and non-SSLP samples only within strata 2-4; thus, the final comparison samples at 7 years of age included 3,558 children/families in 93 SSLP areas and 1,436 children/families in 72 non-SSLP areas.

The demographic characteristics of the final SSLP and non-SSLP samples can be seen in Table 2.2. When strata 2-4 are considered, which are the strata used in SSLP vs. non-SSLP comparisons, there are some demographic differences between the SSLP and non-SSLP samples. One of these revealed greater disadvantage amongst the SSLP sample (i.e., SSLPs had a higher proportion of lone parents), whereas other differences related to ethnicity suggested less disadvantage amongst the SSLP sample (i.e., a higher proportion of white families and a lower proportion of homes where English was an additional language). On other background factors the two samples proved similar (e.g., proportion of mothers having given birth to the target child while under 20 years of age; proportion of households with total incomes below the poverty line). The areas in which SSLP families resided also scored lower on the 2004 overall Index of Multiple Deprivation indicating that they were, on average, more deprived (data not shown).

Table 2.2: Summary of Demographic Characteristics: Imputed data for all seen at age 7

Characteristic	Sure Start (N=5391)							Non-SSLP (N=1436)							probability of NESS-MCS difference for strata 2-4[#] occurring by chance
	Percentage in Each Stratum					In Strata 2-4		Percentage in Each Stratum					In Strata 2-4		
	Strata: SSLP Community Like					Number	%	Strata: SSLP Community Like					Number	%	
	least 1	2	3	4	most 5			least 1	2	3	4	most 5			
Total with imputed data	76	600	1324	1634	1757	3558	66.0		637	471	328		1436	100	
Child's Gender															0.97
Male	46.1	47.2	48.6	50.5	52.5	1752	49.2	-	48.7	50.3	48.8	-	707	49.2	
Female	53.9	52.8	51.4	49.5	47.5	1806	50.8	-	51.3	49.7	51.2	-	729	50.8	
Child's Ethnicity															<0.0001
White	97.4	85.7	85.1	81.6	61.9	2975	83.6	-	81.2	71.5	71.0	-	1087	75.7	
Mixed	1.3	4.2	4.4	4.2	6.2	152	4.3	-	5.2	4.2	4.6	-	68	4.7	
Indian	1.3	1.8	1.2	1.2	1.8	46	1.3	-	3.6	1.9	12.5	-	73	5.1	
Pakistani	-	1.0	2.6	5.3	12.7	127	3.6	-	4.2	11.0	2.7	-	88	6.1	
Bangladeshi	-	0.7	1.1	0.9	6.4	33	0.9	-	0.5	5.7	0.3	-	31	2.2	
Black Caribbean	-	1.8	1.6	1.2	2.2	52	1.5	-	0.8	1.7	2.7	-	22	1.5	
Black Other	-	2.7	1.9	3.7	4.0	102	2.9	-	1.7	1.9	3.0	-	30	2.1	
Other	-	2.2	2.1	1.8	4.8	71	2.0	-	2.8	1.9	3.0	-	37	2.6	
Language in Home															<0.0001
English Home Language	97.4	90.2	90.0	85.7	70.5	3133	88.1	-	87.3	78.6	78.4	-	1183	82.4	
Other Languages	2.6	9.8	10.0	14.3	29.5	425	11.9	-	12.7	21.4	21.6	-	253	17.6	
Maternal Age at Birth of Child															0.66
Not teenage	97.4	95.7	94.1	92.6	91.2	3333	93.7	-	94.5	92.4	93.3	-	1343	93.5	
Teenage (< 20 years)	2.6	4.3	5.9	7.4	8.8	225	6.3	-	5.5	7.6	6.7	-	93	6.5	

[#]Comparing Sure Start and MCS total numbers or means for strata 2-4 with weighting; - Excluded from analysis due to insufficient MCS or SSLP communities

Table 2.2 (continued): Summary of Demographic Characteristics: Imputed data for all seen at age 7

Characteristic	Sure Start (N=5391)							Non-SSLP (N=1436)							Probability of NESS-MCS difference for strata 2-4[#] occurring by chance
	Percentage in Each Stratum					In Strata 2-4		Percentage in Each Stratum					In Strata 2-4		
	Strata: SSLP Community Like					Number	%	Strata: SSLP Community Like					Number	%	
	least 1	2	3	4	most 5			least 1	2	3	4	most 5			
Total with imputed data	76	600	1324	1634	1757	3558	66.0		637	471	328		1436	100	
Maternal Cognitive Difficulties															0.32
No Difficulties Reported	92.1	91.8	90.9	92.6	88.0	3268	91.8	-	90.0	90.2	93.9	-	1306	90.9	
Has Some Difficulties	7.9	8.2	9.1	7.4	12.0	290	8.2	-	10.0	9.8	6.1	-	130	9.1	
Household Deprivation															0.19
Above poverty line[+]	82.9	63.8	55.7	53.7	34.1	1998	56.2	-	68.0	56.3	51.5	-	867	60.4	
Below poverty line[+]	17.1	36.2	44.3	46.3	65.9	1560	43.8	-	32.0	43.7	48.5	-	569	39.6	
Highest Occupation in Household															0.62
Management/Prof.	46.1	30.3	24.7	22.2	14.2	872	24.5	-	34.1	25.1	22.9	-	410	28.6	
Intermediate	15.8	9.2	8.5	10.0	7.1	331	9.3	-	8.5	8.1	8.2	-	119	8.3	
Small Employer	3.9	10.8	6.4	6.2	5.3	251	7.1	-	9.7	10.2	6.4	-	131	9.1	
Lower Supervisory/Tech	10.5	9.2	10.8	9.0	6.8	345	9.7	-	11.5	10.4	11.3	-	159	11.1	
Semi-Routine	9.2	9.2	11.7	13.0	15.0	423	11.9	-	9.9	14.9	13.1	-	176	12.3	
Routine	3.9	8.3	10.1	8.8	8.9	327	9.2	-	8.2	7.0	11.6	-	123	8.6	
Unemployed	10.5	23.0	27.8	30.8	42.7	1009	28.4	-	18.2	24.4	26.5	-	318	22.1	

[#]Comparing Sure Start and MCS total numbers or means for strata 2-4, - Excluded due to insufficient MCS or SSLP communities; [+]Poverty line is £210 per week, 2004-2005

Table 2.2 (continued): Summary of Demographic Characteristics: Imputed data for all seen at age 7

Characteristic	Sure Start (N=5391)							Non-SSLP (N=1436)							probability of NESS-MCS difference for strata 2-4[#] occurring by chance
	Percentage in Each Stratum					In Strata 2-4		Percentage in Each Stratum					In Strata 2-4		
	Strata: SSLP Community Like					Number	%	Strata: SSLP Community Like					Number	%	
	least 1	2	3	4	most 5			least 1	2	3	4	most 5			
Total with imputed data	76	600	1324	1634	1757	3558	66.0		637	471	328		1436	100	
Lone Parent															0.02
Not Lone Parent	86.8	77.8	76.9	71.5	64.5	2654	74.6	-	84.9	81.1	80.2	-	1186	82.6	
Lone Parent	13.2	22.2	23.1	28.5	35.5	904	25.4	-	15.1	18.9	19.8	-	250	17.4	
Work Status Household															0.12
Working Household	89.5	77.0	72.2	69.3	57.3	2550	71.7	-	81.8	76.0	73.5	-	1120	78.0	
Workless Household	10.5	23.0	27.8	30.7	42.7	1008	28.3	-	18.2	24.0	26.5	-	316	22.0	
Highest Education in Household															0.07
Degrees/Higher Education	59.2	33.5	29.3	29.2	20.1	1066	30.0	-	35.2	25.5	26.2	-	430	29.9	
A level	17.1	27.3	27.2	26.8	27.4	962	27.0	-	34.5	37.8	33.8	-	509	35.4	
O level / GCSE	14.5	22.8	24.1	21.7	21.8	811	22.8	-	19.2	21.0	24.7	-	302	21.0	
Other	1.3	7.3	6.9	6.5	6.7	242	6.8	-	3.6	4.0	4.3	-	56	3.9	
None	7.9	9.0	12.5	15.8	23.9	477	13.4	-	7.5	11.7	11.0	-	139	9.7	
Child's Age (Months)															0.16
Mean	87.5	86.3	86.2	86.4	86.4	86.3		-	86.7	86.6	86.5	-	86.6		
SD	3.5	3.2	3.1	3.4	3.3	3.3		-	2.8	2.9	2.9	-	2.9		

[#]Comparing Sure Start and MCS total numbers or means for strata 2-4; - Excluded from analysis due to insufficient MCS or SSLP communities

2.7 Data Collection

The families participating in the NESS longitudinal impact study, the *Study of Children, Families & Services in the Community*, provided extensive information on child and family functioning during the course of a single home visit conducted by a specially trained fieldworker, typically lasting around 90 minutes when children were 9 months of age and then again at 3, 5 and 7 years of age. In the case of home visits to families with 9-month-olds, a survey research workforce under subcontract from the Office of National Statistics carried out data collection. Home visits to families with 3, 5 and 7-year-olds, that involved child assessments as well as parental interviews, were carried out by a field force especially hired and trained for this purpose by the Institute for the Study of Children, Families and Social Issues, Birkbeck, University of London. MCS data were gathered by similar means by several survey research businesses contracted by the MCS team at the Centre for Longitudinal Studies in the Institute of Education, University of London.

During home visits, several sets of data were gathered on child development and family functioning. In addition to these dependent-variable outcome measures, demographic and background information were collected from each family, as well as area characteristics on each community, to serve principally as control variables in the analyses to be presented. Also data on primary school characteristics were obtained from the Office for Standards in Education (Ofsted) and the Department for Education (DfE). Additionally, data on children's Key Stage 1 (KS1) assessment results were obtained from the Department for Education's National Pupil Database. The KS1 data provide assessments of educational achievement that are completely independent of either of the research teams or the parents, and hence are particularly useful for comparison purposes.

The measures delineated below and used in analyses reflect those variables where the procedures within the NESS and MCS studies were sufficiently similar to be comparable across the studies.

2.7.1 Child/Family, Community, School and Study Design Control Variables

A variety of child/family and community variables functioned (principally) as control variables in the analyses to be described (see Appendix B). These included the following:

- *Child Characteristics*: age (in months), age in school year, gender and ethnicity.
- *Demographic, Socioeconomic and Parental Characteristics*: English as only household language (yes, no), maternal age at child's birth (<20 vs. $\geq$ 20), lone parent (yes/no), maternal self-reported cognitive difficulties (some vs. none), household income (below vs. above poverty line), highest individual occupational status in household, highest educational level of household (see Table 2.2), household work status (workless household vs. adult employed).

- *Area characteristics*: Area data, derived from the Index of Multiple Deprivation (IMD, ODPM, 2004) and the 2001 census (for detail see Appendix E), were subject to a principal components analysis that yielded seven area-level factors. For purposes of the current evaluation of SSLP effects, the resulting area-level factor scores function as covariates. The seven area factors were identified as, economic deprivation, large non-Asian ethnic minority present, many children, large Asian/Pakistani population, large transient population with children, large Asian/Bangladeshi population, and large Asian/Indian and student population. In addition the IMD 2004 and an index of urban/rurality were included as area level variables.
- *School characteristics:* At the time of assessment the children would have been attending primary school for about three years. Therefore it is likely that the nature of the school that they attend will affect their development. Hence data on the primary school attended by the child at age 7 years was obtained from Office for Standards in Education (Ofsted) as well as from the Department for Education National Pupil Database. Ofsted produce an overall rating of school effectiveness that was used as a measure of primary school variation and in addition DfE information on school intake, in terms of % of pupils eligible for free school meals (FSM), pupil attendance rate, and the contextualised value added measure (cva) of progress from Key Stage 1 to Key Stage 2 was also used, giving in total four primary school variables as covariates.

2.7.2 Child/Family Dependent/Outcome Variables

The outcome variables for children and families at the 7-year contact and used in analyses are summarised below, with further detail in Appendix C:

- *Academic achievement:* Every child in a state school is assessed at Key Stage 1 (KS1). This occurs towards the end of the school year when the child becomes 7 years old. Typically this will be the child's third year in primary school. Tasks and tests are designed to help inform the final teacher assessment judgement reported for each child at the end of Key Stage 1. Teachers summarise their judgements on children's attainment in relation to the National Curriculum level descriptions. The aim is to reach a rounded judgement that is based on knowledge of how a child has performed over time and across a range of contexts, and takes into account strengths and weaknesses of a child's performance through the key stage. The KS1 assessments used in analysis are reading, writing, mathematics and science. In order to take account of any possible year by year changes in the pattern of KS1 scores, we have used national data on all children in England to create within-year standardised scores for every child. These KS1 standardised scores are not vulnerable to the criticism that differences between SSLP and non-SSLP children may reflect year of measurement effects, in that this strategy eliminates the effect of any year-by-year changes within KS1 data, and provides a fair basis for comparison across samples and across years.

- *Child Physical Health*: Parents were asked about their child's physical health. In

addition child body mass index (BMI) was calculated from height and weight measurements by a researcher.

- *Child social and emotional development:* Forty items dealing with child socio-emotional behaviours were rated by a parent for their child. From these 40 items four underlying factors were extracted by statistical analysis (principal components analysis). These four factors were emotional dysregulation, positive social behaviour, internalising behaviour, self-regulation. These were all obtained by means of parental report.

- *Parental Employment change:* Parental employment was recorded at each stage of the longitudinal study and as improving the employment prospects was one of the aims of Sure Start, change in parental employment status was chosen as an outcome to be investigated. In particular this report examines the likelihood of change in working/workless household status from when the child was 9 months of age to when the child is seven years of age. – Note this is included in our section on change in outcomes and has a different form of analysis to other outcomes in that change since the child was 9 months old is estimated.

- *Parenting and Family Functioning*: Parents were asked a range of questions about their parenting behaviour and the child's activities in the home. Using measures derived from previous research, measures relating to harsh discipline, home chaos, and the home learning environment (HLE) were collected by parent report.

- *Maternal well-being*: life satisfaction; depression; both by parent report.

3. RESULTS

Four stages of analysis are presented:

1. Stage 1 addresses the issue of whether there were across-the-board effects of SSLPs on child and family functioning when children were 7 years of age or in terms of change over time in the case of outcomes measured at both 3 and 7 years of age.
2. Stage 2 seeks to determine whether any effects detected by comparing SSLP and non-SSLP comparison samples in Strata 2-4 may have under, over or accurately estimated overall effects of SSLPs.
3. Stage 3 focuses upon whether detected effects of SSLPs varied across demographically-defined sub-populations (e.g., workless households, lone-parent families) and assesses whether attrition might undermine confidence in any results emerging from the prior analysis.
4. Stage 4 considers whether missing data threatens the confidence in the findings.

It is critically important to appreciate that all analyses outlined in this section were carried out three times:

1. Using only those cases for which there was no missing data at age 7;
2. Using all cases seen at age 7, with any missing data imputed multiple times before being subject to analysis (N=3,558 for SSLP, 1,436 for non-SSLP); and
3. Using all cases seen at age 3 irrespective of whether they were seen at age 7, again with missing data imputed multiple times before being subject to analysis (N=9,192 for SSLP, 1,879 for non-SSLP).

Missing values on all independent and dependent variables were estimated based on standard multiple-imputation procedures in the latter two sets of analyses (Rubin, 1987). The imputation approach represents an attempt to counteract the possibility that cases with missing data differ in some way from cases with complete data and the biasing effects that their exclusion from the analysis could have on the results. Ten imputed data sets were created (for each of the two sets of imputation-related analyses), which ensured that all model estimates will be over 90% efficient. For more detail of the imputation procedure see Appendix D. *Only significant differences between SSLP and comparison areas that emerged in all three sets of analyses are regarded as reliable and meaningful and thus presented and interpreted in this report. This conservative procedure for discerning SSLP effects maximises confidence in the results.* While in the main text only tables showing results which proved significant across all three sets of analyses are presented, parallel tables pertaining to each individual set of results are presented in Appendix F.

3.1 First stage: Overall (across-the-board) Effects of SSLPs

The first stage of data analysis was designed to assess the main (or across-the-board) effects of SSLPs on each dependent variable, after taking into account pre-existing

differences between SSLP and comparison families, communities and schools in their demographic characteristics (shown in Table 2.2). Thus, an overall main effect involves a significant difference between SSLP and comparison communities on an outcome without taking into consideration the possibility that sub-populations might be differentially affected by SSLPs. In other words, it addresses the question as to whether, on average across all types of children and families, effects of SSLPs emerged. (Note that these across-the-board comparisons do not involve the entire SSLP sample, but only those areas in strata 2-4, the ones most similar across SSLP and non-SSLP samples).

In order to determine whether main effects of SSLPs on child development and family functioning were detectable, the data were analysed using multilevel models, which take into account the hierarchical structure of the data, with children and families nested within communities, some of which are SSLP communities and some comparison communities. Linear models are used for the continuous measures and logistic models for binary outcomes. Summary statistics (i.e., means, standard deviations) for the SSLP vs. non-SSLP comparisons using complete data for all cases seen at 7 years, are presented for each of the outcomes in Table 3.1. The overall results of SSLP vs. non-SSLP comparisons for the analyses (i.e. estimated effects) can be seen in Table 3.2. *The effect sizes listed in Table 3.2 are the average of effect sizes across all sets of analyses, in terms of differences between groups measured in standard deviation units*. Effect sizes represent the magnitude of effects in a way that allows comparison between different variables regardless of their scale of measurement. It also affords a convenient way of comparing effects detected in the evaluation of one intervention project with those of another.

Statistical significance was used as an indication of the confidence that the finding is genuine rather than occurring by chance. However, the fact that a given finding is statistically significant does not necessarily mean that it is important. A finding of a small difference between two conditions may well be statistically significant but may be trivial, particularly if it applies to only a small group. Therefore a measure of the size of the difference is crucial. Effect size is the name given to statistical measures of the magnitude of a difference. Effect sizes can help us understand the depth and importance of the differences observed. The most common measure of effect size used is the difference between two scores divided by the standard deviation. In comparing SSLP and non-SSLP samples the effect size was calculated by dividing the difference in scores by the standard deviation of the total sample. In studies using RCTs (e.g. drug trials) effect sizes of around 0.2 are usually regarded as ‘small’, of 0.5 as ‘medium’ and 0.8 or greater as ‘large’. However the importance for policy of a particular effect size is also related to the proportion of the population affected, and the effects associated with factors linked to policy are typically in the “small” range, e.g., the effect size for eligibility for free school meals (FSM) is 0.23 for literacy and 0.15 for numeracy (Key Stage 2 results, Sammons et al., 2008). Where population-scale differences are involved smaller effect sizes may well be of policy significance. For example a population-wide change for IQ with an effect size as small as 0.1 would have profound implications for education and the economy.

Table 3.1 Summary of Outcome Measures: Complete data for all seen at age 7

Outcome Measures	Sure Start						Non-Sure Start					
	Summary for Each Stratum					For Strata 2-4	Summary for Each Stratum					For Strata 2-4
	Stratum: SSLP Community Like						Stratum					
	least				most		least				most	
	1	2	3	4	5		1	2	3	4	5	
Physical Health												
BMI (standardised) Mean	0.53	0.50	0.55	0.63	0.58	0.58	-	0.49	0.32	0.50	-	0.44
SD	1.00	1.23	1.18	1.25	1.30	1.20		1.19	1.09	1.29		1.19
General health	4.49	4.39	4.38	4.42	4.30	4.39	-	4.38	4.44	4.46	-	4.42
SD	0.72	0.84	0.82	0.82	0.91	0.82		0.81	0.81	0.85		0.82
Child Educational Development												
Key stage 1 Reading Mean	0.14	-0.04	0.00	-0.06	-0.15	-0.03	-	0.02	0.02	-0.02	-	0.01
SD	1.01	0.94	0.99	0.94	0.95	0.94		0.91	0.85	0.98		0.90
Key stage 1 Writing Mean	0.13	0.02	-0.01	-0.10	-0.16	-0.05	-	-0.03	0.01	-0.05	-	-0.02
SD	0.99	0.93	0.94	0.92	0.94	0.93		0.91	0.85	0.91		0.88
Key stage 1 Maths Mean	0.06	-0.04	0.03	-0.13	-0.18	-0.06	-	-0.04	-0.00	-0.05	-	-0.03
SD	1.09	0.96	0.92	0.93	0.98	0.93		0.95	0.89	0.95		0.93
Key stage 1 Science Mean	-0.01	-0.09	-0.07	-0.11	-0.23	-0.09	-	-0.03	-0.06	-0.09	-	-0.05
SD	0.96	0.96	0.90	0.86	0.87	0.89		0.97	0.82	0.86		0.90
Child Behaviour and Social Development												
Emotional dysregulation Mean	1.52	1.60	1.64	1.65	1.69	1.64	-	1.62	1.67	1.64	-	1.64
SD	0.40	0.42	0.45	0.46	0.45	0.45		0.47	0.47	0.47		0.47
Positive social behaviour Mean	2.72	2.75	2.72	2.71	2.69	2.71	-	2.72	2.67	2.70	-	2.70
SD	0.28	0.28	0.30	0.30	0.32	0.30		0.27	0.28	0.30		0.28
Internalisation Mean	1.26	1.34	1.31	1.35	1.36	1.34	-	1.35	1.35	1.35	-	1.35
SD	0.27	0.36	0.32	0.34	0.35	0.34		0.33	0.32	0.35		0.33
Self regulation Mean	2.46	2.45	2.43	2.44	2.42	2.44	-	2.44	2.41	2.45	-	2.43
SD	0.34	0.37	0.38	0.38	0.38	0.37		0.37	0.36	0.37		0.36

Outcome Measures	Sure Start						Non-Sure Start					
	Summary for Each Stratum					For Strata 2-4	Summary for Each Stratum					For Strata 2-4
	Stratum: SSLP Community Like						Stratum					
	least				most		least				most	
	1	2	3	4	5		1	2	3	4	5	
Maternal Wellbeing												
Mother's self rated depression Mean	1.55	1.64	1.70	1.69	1.79	1.68	-	1.58	1.61	1.64	-	1.60
SD	0.65	0.77	0.83	0.79	0.90	0.80		0.68	0.70	0.81		0.72
Mother's satisfaction with life Mean	7.91	7.61	7.47	7.47	7.48	7.49	-	7.35	7.28	7.31	-	7.32
SD	2.02	2.08	2.14	2.15	2.30	2.13		1.83	1.95	2.03		1.91
Parent and Family Functioning												
Harsh discipline in the home Mean	2.51	2.56	2.57	2.57	2.55	2.57	-	2.67	2.69	2.60	-	2.66
SD	0.45	0.51	0.51	0.51	0.54	0.52		0.57	0.57	0.68		0.60
Chaos in the home Mean	2.06	2.09	2.17	2.18	2.15	2.16	-	2.40	2.49	2.35	-	2.42
SD	0.72	0.91	0.90	0.85	0.85	0.88		1.06	1.09	1.17		1.10
Home learning environment Mean	3.73	3.76	3.75	3.74	3.77	3.75	-	3.66	3.67	3.79	-	3.25
SD	0.59	0.66	0.73	0.71	0.76	0.70		0.66	0.68	0.61		0.87

- Excluded from analysis due to insufficient MCS or SSLP communities; SD=Standard Deviation

After taking into consideration pre-existing family and area background characteristics (in all sets of analyses), across all three sets of analyses comparing children and families living in SSLP areas and those living in similar areas not receiving SSLPs, there were three positive effects of SSLPs and twelve non-effects with respect to child and family functioning when children were 7 years old:

For maternal well being and family functioning, in comparison to mothers in non-SSLP areas, mothers residing in SSLP areas reported:

- *providing a more cognitively stimulating home learning environment.*
- *providing a less chaotic home environment for their children.*
- *engaging in less harsh discipline of their children.*

There was no consistent effect on measures of maternal well-being. Also there were no consistent differences on any of the four child educational development outcomes, or on four child social and behavioural outcomes or on the two child health outcomes.

In summary, across a total of 15 outcomes[4] evaluated, significant effects of SSLPs emerged in the case of 3 outcomes. The estimated SSLP effects are displayed in Table 3.2 for the results from the analysis of imputed data for all cases seen at 7 years. The data in the table chronicle the degree of difference, including confidence intervals, between SSLP and comparison areas, after adjusting for child and family background factors, community characteristics and school characteristics on measurements made at age 7 (shown in Table 2.2). Positive values on the mean difference scores indicate that SSLP areas scored higher than the comparison areas, as do odds ratios greater than 1.00; negative values of mean difference scores indicate the opposite, as do odds ratios less than 1.00. The effects of the strata, demographic, family characteristics and area level measures are shown in Appendix F.

Although Table 3.2 presents the main effects of SSLPs, that is, whether the SSLP and comparison samples differed significantly across the entire sample, net of control variables, on age 7 measurements, these results do not indicate whether effects of SSLPs varied for specific subgroups related to gender, worklessness or income. Such interaction effects are reported in the third major stage of analysis. It is possible that a result does not accurately reflect the situation—due to subsequent qualification by an interaction effect; caution is therefore warranted in drawing conclusions on the basis of these main-effect results alone. An interaction effect would indicate that the effect of exposure to SSLPs was not the same across different subgroups. One might discover, for instance, that SSLPs affected the parenting of lone parents more than those in two parent families.

[4] Definitions of the outcomes can be found in Appendix C.

Table 3.2: Estimated Effects of Sure Start at 7 years- averaged across analyses

Outcome Measures	Sure Start Main Effects[#]		
	Mean difference in sd units Effect size	**95% CI**	**p**
Child Behaviour and Social Development			
Emotional dysregulation	-0.04	-0.14 to 0.06	0.29
Positive social behaviour	0.09	-0.01 to 0.17	0.14
Internalising	-0.08	-0.19 to 0.03	0.19
Self regulation	0.01	-0.08 to 0.10	0.54
Child Physical Health			
BMI	0.05	-0.06 to 0.16	0.38
Physical health	0.00	-0.08 to 0.09	0.37
Child Educational Development			
Key stage 1 Reading attainment	0.02	-0.08 to 0.12	0.71
Key stage 1 Writing attainment	0.03	-0.07 to 0.14	0.56
Key stage 1 Maths attainment	0.05	-0.06 to 0.15	0.37
Key stage 1 Science attainment	0.03	-0.08 to 0.15	0.61
Maternal Wellbeing			
Mother's life satisfaction	0.06	0.00 to 0.12	0.14
Mother's self rated depression	0.09	0.02 to 0.17	0.08
Parent and Family Functioning			
Harsh discipline in home	**-0.17**	-0.26 to -0.08	**0.00**
Chaos in home	**-0.29**	-0.43 to -0.14	**0.00**
Home learning environment	**0.66**	0.53 to 0.77	**<0.0001**

[#] Effects are adjusted for child, family and area characteristics and strata

Change from 3 to 7 years: Eight of the dependent measures included in the 7-year analyses were also measured in similar fashion at the 3-year assessment, so it was possible to evaluate the effects of SSLPs on *change* in functioning from 3 to 7 years of age. The same statistical models used to produce the results just summarised were rerun, with one modification: the 3-year version of the 7-year outcome to be predicted was added as a covariate, thereby changing the outcome from a measure of functioning at 7 years of age to one of *change* from 3 to 7 years. The exception to this concerns change in workless household status, which is considered as change since the child was 9 months old. Results of this analysis are presented in Table 3.3. *The effect sizes quoted in Table 3.3 are the average of the effect sizes across all three sets of analyses*.

Table 3.3: Estimated Effects of Sure Start for change between 3 and 7 years – averaged across analyses

Outcome Measures	Sure Start Main Effects[#] Estimated Effect		
	Mean difference in sd units Effect size	**95% CI**	**p**
Child Behaviour and Social Development			
Emotional dysregulation	-0.01	-0.09 to 0.08	0.44
Positive social behaviour	0.01	-0.08 to 0.09	0.50
Internalising	-0.03	-0.12 to 0.05	0.51
Self regulation	-0.05	-0.14 to 0.05	0.36
Maternal Wellbeing			
Mother's life satisfaction	0.06	-0.01 to 0.13	0.33
Mother's self rated depression	-0.05	-0.14 to 0.04	0.19
Parent and Family Functioning			
Harsh discipline in home	**0.10**	-0.18 to -0.01	**0.03**
Home learning environment	**0.58**	0.46 to 0.69	**<0.0001**

[#] **Effects are adjusted for child, family and area characteristics and strata**

CI=Confidence interval

Inspection of the table reveals that of eight repeatedly measured dependent variables, two showed evidence of positive SSLP effects on change over time in child or parent functioning in all three sets of analyses.

In comparison with those in non-SSLP areas, mothers in SSLP areas:

- *Reported more positive change in the home learning environment (i.e., greater improvement),*
- *Reported more positive change in harsh discipline (i.e., greater decrease).*

There were no differences associated with SSLPs on change from age 3 to 7 years in child outcomes, nor was there any difference in change in workless household status from 9 months to 7 years (see Appendix C for explanation of the measures).

3.1.1 Summary of main effects

Across 15 7-year outcomes evaluated, significant main or across-the-board effects of SSLPs emerged in the case of three outcomes. After taking into consideration pre-existing family, area (community), and school characteristics, the three sets of analyses comparing children and families living in SSLP areas and those living in similar areas not receiving SSLPs revealed three apparently positive effects of SSLPs. Mothers in SSLP areas reported engaging in less harsh discipline and providing a less chaotic and more cognitively stimulating home learning environment than counterparts from the MCS not living in SSLP areas.

With regard to change from age 3 to age 7 years, of eight repeatedly measured dependent variables, two indicated SSLP effects on change over time in child or parent functioning in all three sets of analyses. Mothers in SSLP areas reported more positive change in two outcomes relative to those residing in comparison areas in the use of harsh discipline (i.e., greater decrease), while evincing more positive change as well in the home learning environment (i.e. greater improvement). The analyses of change in workless household status from 9 months to 7 years revealed no difference between SSLP and non-SSLP areas.

3.2 Second Stage: Did first stage analysis over/underestimate SSLP effects?

The across-the-board SSLP vs. non-SSLP comparisons just summarised did not include substantial numbers of SSLP children/families—those from Strata 1 and 5—because there were insufficient numbers of children/families in the non-SSLP sample in stratum 5 or in the SSLP sample in stratum 1 to afford reliable comparisons. To determine whether the main effects of SSLPs detected and reported in the preceding section *might* have been different had it proven possible to include all SSLP children/families a second stage of analysis was conducted comparing all SSLP children and families from stratum 2-5 with one another on the outcome variables (after adjusting for all covariates); because Stratum 1 included very few SSLP children/families it was excluded from this analysis. If the *within-SSLP comparisons* involving stratum 2-5 revealed differential functioning of the SSLP children/families not included in the SSLP vs. non-SSLP comparison, this would suggest that the SSLP vs. non-SSLP comparisons just reported might have either overestimated or underestimated SSLP effects.

The results of this second stage of analysis—across all three data sets (i.e. complete, imputed 1 & 2) —revealed that SSLP children/families from Stratum 5 did not differ on any dependent measures from those in Strata 2-4 after adjusting for background characteristics. This suggests that had it been possible to include Stratum 5 children in the analysis reported in the preceding section the results would not have changed and thus those SSLP main effects discerned in the preceding analysis are generalisable to Stratum 5 children/families. In summary the SSLP effects detected appear to apply equivalently across all SSLPs regardless of level of deprivation. Clearly, this is an *argument by inference* and not an indisputable conclusion.

3.3 Third Stage: Differential Effects of SSLPs on Specific Sub-populations

Having detected indications of some positive effects of SSLPs on family functioning and maternal well-being when children were 7 years of age and in change from 3 to 7 years, as well as having found that these appear to apply to all SSLP areas (i.e., even the most deprived areas excluded from the main analysis), the third stage of analysis was designed to determine whether effects of SSLPs were the same across various population sub-groups. This issue is particularly important given early findings from the initial cross-sectional study of 3-year-olds showing that

various demographically-defined sub-populations were differentially affected by SSLPs (NESS, 2005; Belsky et al., 2006).
In order to examine whether SSLP effects varied across select sub-populations, five demographic variables were chosen, because of their policy relevance, to address the issue of sub-population-specific effects of SSLPs. More specifically, 2-way interactions involving SSLP status and each of the following factors were tested for each outcome measure after controlling for the child, family and area characteristics (again using only Strata 2-4 children/families):

- child gender
- teenage parenthood (i.e. <20 years at delivery)
- lone parenthood (i.e. no partner living in home)
- workless household (i.e. no adult employed in home), and
- household deprivation (i.e. <210 p.w., >£210 p.w. or below poverty line)

(Note: The £210 p.w. income figure was chosen as this was 60% of the median income for the country at the time of the 9-month data collection, and people with incomes below this figure were officially regarded as poor.)

Table 3.4: Summary of interaction effects between SSLP status and select demographic variables that were replicated in all three data sets

Outcome Measure	Child's Gender	Maternal Age at Birth of Child	Poverty	Lone Parent	Work Status Household
Child Physical Health					
General Health					
Standardised BMI					
Child Educational Development					
Key stage 1 Reading attainment					
Key stage 1 Writing attainment					
Key stage 1 Maths attainment					
Key stage 1 Science attainment					
Child Behaviour and Social Development					
Emotional dysregulation					
Internalisation					
Positive social behaviour					
Self regulation					
Maternal Wellbeing					
Mother's self rated depression					
Mother's satisfaction with life				*	*
Parent and Family Functioning					
Harsh discipline in the home					
Chaos in the home	*				
Home learning environment					

* Statistically significant in all 3 data sets

Inspection of Table 3.4 indicates that there were 3 significant interactions involving SSLP status and demographic factors in the prediction of 7-year outcomes that replicated across all three data sets. In that a total of 75 interactions were tested, this number of significant interactions is less than might be expected by chance. However the same outcome is involved in 2 out of the 3 significant interactions, indicating that this may not be a chance result, when also looked at in terms of change over time.

Table 3.5: Summary of interaction effects between SSLP status and select demographic variables for models of change 3 to 7 years that were replicated in all three data sets

Outcome Measure	Child's Gender	Maternal Age at Birth of Child	Poverty	Lone Parent	Work Status Household
Child Behaviour and Social Development					
Emotional dysregulation					
Internalisation					
Positive social behaviour					
Self regulation					
Maternal Wellbeing					
Mother's self rated depression					
Mother's satisfaction with life				*	*
Parent and Family Functioning					
Harsh discipline in the home					
Home learning environment					

* Statistically significant in all three data sets

For the analysis of change from 3 to 7 years in outcomes measured repeatedly (see Table 3.5), there were 2 significant interactions —out of 40 evaluated--involving SSLP status and demographic factors replicated across all three data sets. This is again what might be expected by chance, but as they involve the same outcome as the analysis presented in Table 3.4 this suggests that they are not chance results. Details of the interaction analyses are reported in Appendix G.

In summary while the significant results from analyses of interactions for both 7 year outcomes and change from 3 to 7 years were less than might be expected by chance, the same outcomes were involved, suggesting that these were not chance results. Hence it was decided to investigate all these interactions, whether for outcomes or change from 3 to 7 years, further.

Further analyses for boys and girls separately revealed that the gender interactions related to the outcome home chaos resulted from this outcome being significantly lower for boys in SSLP areas than in comparison areas, while for girls there was an effect in the same direction that only approached statistical significance.

For the outcome mother's satisfaction with life, separate analyses for subgroups revealed that the interactions linked to lone parenthood were the result of lone mothers reporting greater life satisfaction in SSLP areas than in comparison areas, while for mothers in two- parent households there were no differences in life satisfaction associated with living in a SSLP area. In addition with regard to the significant interactions involving mother's satisfaction with life and workless/working household, the significant interactions resulted from mothers in workless households reporting greater life satisfaction in SSLP areas than in comparison areas, while for mothers in working households there were no differences in life satisfaction

associated with living in a SSLP area. These patterns of interactions held for both the analysis of the 7 year outcome and change from 3 to 7 years.

3.4 Fourth Stage: Threats to confidence in detected SSLP effects

Given the circumstances under which the NESS impact study was carried out, the non-SSLP (MCS) element of which was not under the control of the NESS impact study team, there are certain threats to the confidence that can be placed in the SSLP effects reported. Significantly, because there was selective attrition within each of the samples studied, SSLP and non-SSLP, the possibility exists that the results are biased as a result of this.

In both samples some children seen at 3 years were not seen again at age 7. In light of this, it is reasonable to ask whether the SSLP effects detected would generalise to children/families not seen at age 7. To address this issue, the SSLP children and families seen at 3 years but not at 7 years were compared on a range of demographic variables with those seen on both occasions, *separately by strata (i.e. attrition group vs. strata1; attrition group vs. strata 2 etc.)* (see Appendix B). If the SSLP children/families not seen at age 7 proved more deprived than those seen at both ages of measurement—and, especially, more deprived than those in stratum 5--this would undermine confidence in the interpretation attributed to the discovery reported in the second stage of analyses that stratum 5 children/families were not different from those in strata 2-4. It was this *absence of difference* that provided the basis for the earlier observation that the SSLP effects detected in the first stage of inquiry—which only involved strata 2-4—probably generalised to stratum 5 children/families, that is, those most likely to be in SSLP areas.

The attrition analysis revealed that risks arising from selective attrition appear not to seriously threaten the confidence to be placed in the SSLP effects detected in the first stage of analysis (*and, which were found, in the second stage of analysis, to be probably generalisable to stratum 5 children).* This is because on some measures one sample appeared more disadvantaged, whereas on other measures the reverse was the case. Specifically, although the SSLP sub-sample not seen at age 7 (attrition group) proved to be significantly *less advantaged* than those in strata 1 to 4, seen at both 3 and 7 years, of the eight variables that indicate a significant difference between the sample seen at 7 years of age and those not seen at 7 years of age, six variables (i.e. ethnicity, English first language, poverty, occupational status, lone parent and workless household) indicate that the sample seen at 7 years in stratum 5 are *less* disadvantaged, and only two variables (i.e. teen mother and, parent education) indicate that the sample seen at 7 years in stratum 5 are *more* disadvantaged than the attrition group (not seen at 7 years of age). This indicates that, despite attrition, the profile of people lost to the study falls within the range of people still in the study so, despite losing them, the results are still generalisable to the SSLP population as a whole. In other words, had the attrition group also been followed up at age 7 the absence of differences on test outcomes between stratum 5 and the other strata would have remained. Once again this would lead to the conclusion that the results from the first stage of analysis involving only strata 2-4 sub-samples appear generalisable to stratum-5 children/families, those most likely to be in SSLP areas.

4. SUMMARY OF RESULTS

After taking into consideration pre-existing family, area and school characteristics four positive effects of SSLPs emerged from 15 outcomes at age 7, two of which applied to the whole population and two of which applied to sub-populations. For the whole population, mothers in SSLP areas, relative to counterparts not living in SSLP areas reported:

(1) engaging in less harsh discipline;

(2) providing a more stimulating home learning environment for their children;

additionally for sub-populations, mothers in SSLP areas reported:

(3) providing a less chaotic home environment for boys (not significant for girls);

(4) having better life satisfaction (lone parent and workless households only).

Note that results for sub-populations can be as important as those for the total population, and knowing about sub-population differences can inform the targeting of services. This is increasingly important as children's centre services are increasingly targeted at the most vulnerable, and also service delivery may be targeted differently for specific sub-populations.

Examination of change over time in child and family function as a function of living in an SSLP area, yielded positive SSLP effects on three of eight repeatedly-measured outcomes when the focus of evaluation was on change in parent and child functioning between 3 and 7 years (or between 9 months and 7 years for workless household status). Mothers in SSLP areas relative to those in comparison areas:

(1) showed a greater improvement in the home learning environment;

(2) reported a greater decrease in harsh discipline;

additionally for sub-populations, mothers in lone parent and workless households in SSLP areas reported:

(3) greater improvement in life satisfaction.

Finally, there were no differences associated with SSLPs for change in child outcomes.

Note that there were only beneficial effects and no negative effects discerned in any of the analyses.

The effects of SSLPs appeared to be the same in the most deprived SSLP areas relative to those somewhat less deprived (but still deprived) areas. Hence it is concluded that the evidence is compatible with the view that SSLPs had similar effects across the demographic spectrum of children and families served by SSLPs. This together with the SSLP effects discerned for lone parent and workless families indicates that SSLPs are likely to be reaching all sections of the populations served.

5. CONCLUSIONS

What then has emerged from this 7-year-old phase of the NESS longitudinal study?

Detected "effects" of SSLPs on Children/Families
To take account of missing data effectively the decision was made to evaluate SSLP effects in three different ways. One set of analyses only included those seen at age 7 with complete measurement data and two other sets of analyses were based on multiple imputation of missing data. One involved imputing missing data on any cases seen at age 7, whereas the other involved imputing missing data on any cases seen at age 3. Evaluation of SSLP effects involved child/family functioning at age 7 for 15 outcome measures, *change from age 3-7* in child/family functioning for eight repeatedly measured outcomes and *change from 9 months to 7 years* in household workless status. Finally, detected effects of SSLPs were only judged to be meaningful and worthy of consideration if they proved significant across all three sets of analyses.

After taking into consideration pre-existing family, area and school characteristics four positive effects of SSLPs emerged from 15 outcomes at age 7, two of which applied to the whole population and two of which applied to sub-populations. For the whole population, mothers in SSLP areas, relative to the comparison group, reported:

(1) engaging in less harsh discipline;
(2) providing a more cognitively stimulating home learning environment for their children;

additionally for sub-populations, mothers in SSLP areas reported:

(3) providing a less chaotic home environment for boys (not significant for girls);
(4) having better life satisfaction (lone parent and workless households only).

Change over time
Additional evidence of positive SSLP effects for the whole population emerged on three of eight repeatedly-measured outcomes when the focus of evaluation was on change in parent and child functioning between 3 and 7 years. Mothers in SSLP areas relative to those residing in comparison areas:

(1) showed a greater improvement in the home learning environment;
(2) reported a greater decrease in harsh discipline;

additionally for sub-populations, mothers in SSLP areas reported:

(3) greater improvement in life satisfaction (lone parent and workless households only).

These latter effects for lone parent and workless households can be regarded as evidence that SSLPs are being successful in affecting 'hard to reach' groups within SSLP areas, which in the early stages of SSLPs had appeared to be a problem.

It is noteworthy that there were only beneficial effects of SSLPs and no negative effects discerned in any of the analyses in this report. Also, the discerned SSLP effects appear to apply (1) to all areas served by SSLPS regardless of level of deprivation, and (2) to all children and families regardless of family deprivation. This indicates that since the earliest stages of the evaluation SSLPs have moved forward and now appear to be engaging the most vulnerable groups in the most deprived

areas. This suggests that SSLPS really did respond to the early findings, and remedied the problems with engaging vulnerable groups.

One potential problem with the evidence in this report is that the positive SSLP effects reported all derive from parental report. This limitation raises the possibility that the effects reflect differential response bias in SSLP areas possibly related to mothers greater use of services and contact with SSLP staff who may emphasise certain behaviours as being more valued. However if this were so then the differential response of SSLP mothers should apply across the population within SSLP areas. However it was found that significant reduced home chaos was reported more often by mothers of boys. It is difficult to think why mothers of boys should be more affected by reporting bias than mothers of girls. Hence this finding may be regarded as evidence against the interpretation that the positive SSLP results reported here reflect reporting bias.

When the children were 5 years of age it was found that there had been a greater reduction in workless household status from 9 months to 5 years in SSLP areas than in comparison areas. However when the comparison was undertaken for change in workless household status from 9 months to 7 years there was no SSLP effect. This suggests something is happening between 5 and 7 years. Analysis of change in workless household status between the child being 5 and 7 years revealed a significantly greater increase in workless household status for SSLP areas. In that the SSLP area data were collected between 2009 and 2011 while the non-SSLP area data were collected two years earlier this suggests the possibility of a timing effect. This would be consistent with parents in the study in SSLP areas experiencing worse unemployment effects than their counterparts as a consequence of the economic recession that started to accelerate from 2008 onwards, i.e. having greater impact on the SSLP sample. This is an example of the time of measurement problems mentioned earlier.

There were only beneficial effects of SSLPs in this 7-year-assessment phase, making this stage of inquiry similar to the earlier phase in which only positive effects emerged when children were 3 years of age (NESS, 2008; Melhuish et al., 2008a). Also the discerned effects appeared to be applicable to all SSLP areas, but some effects only apply to some parts of the population in SSLP areas. The results discerned in this 7-year-old phase of the NESS impact study provide some support for the view that government's effort to support children/families via the original area-based approach to Sure Start paid off, although only with respect to family functioning and maternal well-being, as there were no consistent effects for child outcomes at 7 years of age. Nevertheless, the fact that there were no effects on child functioning raises questions about "return on investment". This is particularly the case with regard to children's capabilities to succeed in school as a result of academic and social skills, at least as measured by KS1 scores.

It might be thought that SSLP effects upon parenting would feed through to more detectable improvements in child outcomes. Although there were some positive effects for child social development at 3 years, no such evidence emerged at age 5 or age 7. While there are significant positive effect sizes for parenting, they were relatively small in the early stages of this longitudinal study, and evidence from the EPPE study indicates that it is the home learning environment (HLE) at age 3, and

not later, which is predictive of later child outcomes (Melhuish et al., 2008c; Melhuish, 2010), and thus any consequent effect upon child development may simply be too small to detect. Also as the effects of HLE appear to be particularly important in the first 3 years, a time when language development is most rapid and most open to environmental influence, it may be that the SSLP effects on HLE are happening too late to improve language development, which itself underpins later cognitive and social development. This explanation fits with the pattern of results seen over the 7 years of the longitudinal study. In addition any effects for child outcomes may be overwhelmed by children's pre-school and/or primary school experiences, which were equivalent for SSLP and non-SSLP groups.

Looking across the longitudinal study some patterns have emerged. By age 3 there were beneficial effects for family functioning and maternal well-being. Similar effects also occurred at 5 and 7 years. For the children at age 3 there appeared to be some beneficial effects of SSLPs in terms of social development but these had disappeared by age 5 and age 7, probably because almost all children were receiving pre-school education between ages 3 and 5, and primary school experience from 5 to 7 years. At age 3 the average child in SSLP areas was functioning at only about the 15th percentile (NESS, 2005) and at ages 3, 5 and 7 there were no SSLP effects related to cognitive or linguistic development. It would appear likely that if SSLPs and related programmes are to achieve the desired effect upon cognitive-linguistic development and related later educational attainment then it will be necessary to either produce improvements in cognitive-linguistic development at age 3 or boost such development subsequently (preferably both). In this regard producing improvements in basic language development is absolutely crucial.

There is evidence that high quality childcare between the ages of 0-3, where high quality means providing a high level of responsive interaction, can benefit language development both for the general population (Melhuish et al., 1990) and for disadvantaged groups (Campbell, & Ramey, 1994). In addition the strongest evidence for population-wide early years programmes affecting child development concerns high quality pre-school education being associated with improved cognitive and social development (Melhuish et al., 2008b; Sylva et al., 2010; Vandell et al., 2010; Melhuish et al., 2010b; Melhuish 2011). Even greater effects are found for disadvantaged groups (Schweinhart, et al., 1993; Magnuson, Ruhm & Waldfogel, 2007). While pre-school education is part of what SSLPs (now children's centres) offer, it has also been available to children in non-SSLP areas. Since April 2004 all 3 and 4 year olds have been entitled to free part-time (15 hours/week) pre-school provision and 95% of the eligible population in England take up this offer (DfE, 2011). Hence there are unlikely to be differences in pre-school education experiences, for which evidence related to child development is strongest, between children in or not in SSLP areas. This is particularly likely to be the case as the quality of pre-school provision in SSLP areas was found by NESS to be similar to that available nationally (Melhuish et al., 2010a). This equivalence of pre-school education experience could be responsible for the failure to detect SSLP effects on children at age 5 and age 7. That is, it could be the case that developmental advantages associated with SSLPs at age 3 prove non-existent at age 7 because by this time all children were being exposed to pre-school (and also primary school) provision, which results in "catch up" for those children in non-SSLP areas. Importantly, SSLPs have been found to be immensely popular with parents (see user

satisfaction survey results from Local Programme Evaluation findings[5]), and as demonstrated by evidence in this report, are reaching the parents who are likely to be very poor. This established infrastructure needs to give greater emphasis to services that will improve child outcomes, particularly language development, if child outcomes are to be enhanced for the children served. In relation to this last point a previous NESS report (Melhuish et al., 2010a) explored the quality of pre-school provision in SSLP areas and found that attending higher quality pre-school was associated with better language development.

Shifting the curve or lifting the "tail"

Early intervention strategies aimed at improving child outcomes can be considered in terms of whether they shift the population curve for an outcome (i.e. the population generally benefits) or whether they lift the "tail" of the population (i.e. benefit is for the lower extreme of the population). The services offered by SSLPs (now children's centres) consisted largely of various parental support programmes plus early childhood education and childcare (ECEC). In looking at the evidence on early interventions, there are several examples of intensive home-visiting parenting programmes producing benefits for small highly targeted disadvantaged groups (for a review see Schrader-McMillan, Barnes, & Barlow, 2012, in press), and Barrett (2010) provides a review of what works best in parenting programmes. In addition there are several examples of ECEC producing benefits for disadvantaged groups (Barnett, 2011).

For whole populations, the only strong evidence for shifting the population curve is for provision of high quality ECEC, particularly early childhood education (Melhuish, 2011; Sylva et al., 2010). In the case of this study the children in SSLP areas have been, and continue to be, eligible for a range of parenting support plus 15 hours of ECEC through the free entitlement. The children in non-SSLP areas are eligible for 15 hours/week of ECEC through the free offer for 3 and 4-year-olds as mentioned earlier. In all cases the take-up of ECEC from 3 years on is around 95%. Also the quality of ECEC in SSLP areas is similar to that in England as a whole (Melhuish et al., 2010a). Hence the experience of ECEC from 3 years onwards both in amount and in quality is similar for both SSLP and non-SSLP areas.

In a longitudinal study of 2000+ children in deprived areas of Chicago, Reynolds et al., (2011) examined which components of a comprehensive early childhood intervention programme, which included parenting, ECEC and other interventions, were associated with long-term developmental benefits for children. They found that only ECEC (which was of high quality in their study) was independently associated with long-term benefits in terms of educational attainment, social adjustment and occupational success. Therefore, the results of this study, demonstrating no overall improvement for children's outcomes as a result of living in SSLP areas at age 7, are consistent with the evidence that only high quality ECEC produces population-wide benefits. One way of summarising this evidence is that only high quality ECEC will shift the population curve, while intensive parenting programmes may have some benefit for the targeted "tail" of the population. This interpretation is consistent with the evidence that high quality early childhood education and care (ECEC) (e.g.

[5] http://www.ness.bbk.ac.uk/support/local-evaluation-findings

Melhuish, 2011) will shift the population curve for child outcomes, and this is the only type of early intervention for which evidence is currently available for shifting the population curve through enhancing the development of all children in the relevant population, rather than lifting the "tail" of the population through targeted intervention. Other types of intervention strategy have been found to be successful in helping targeted small groups of families with young children, notably well-implemented "manualised" parent support programmes (Barrett, 2010), and such programmes are sometimes used by Sure Start programmes, and they can help lift the tail of the population. Also there are examples of childcare in the first 3 years combined with parent support (e.g., Early Head Start; Love et al., 2002) that have also been found to have substantial impact for disadvantaged populations.

Methodological Constraints

In interpreting these findings some important constraints are relevant. Under ideal scientific circumstances areas would have been randomly assigned to receive an SSLP or not. This would generate the strongest policy-related conclusions. It also would have been beneficial to have data on children/families in SSLP and non-SSLP areas collected by the same research team and at the same time. The fact that this was not the case weakened the NESS impact study in ways beyond the control of the NESS team. In particular the two year gap between the measurements in the SSLP and non-SSLP samples meant that, with the exception of the child educational (KS1) data, any and all discerned effects of SSLPs—whether positive or negative in character- potentially could be attributed to time-related changes in communities and/or the larger society that had nothing to do with SSLPs. These constraints highlight the importance of giving early consideration, during the planning phase of an evaluation, to the trade-offs involved in compromising on fundamental evaluation design issues.

Final comments

Sure Start has undergone evolutionary change since its inception in 1999. To some extent evaluation results have influenced this process. The early results indicated that lack of specification of how goals were to be achieved in service delivery led to great programme diversity, which may have meant that many SSLPs were not delivering services likely to improve child outcomes. Later policy changes have led to tighter guidelines for service delivery and also staff themselves have developed and become better trained and more proficient. However, there is still scope for further development.

The contrast between the first cross-sectional and subsequent NESS longitudinal findings suggests that children and families in the longitudinal phase had greater exposure to SSLPs/children's centres that had more time to 'bed-in' and develop. This contrast between the first cross-sectional phase and subsequent findings indicates that such early interventions can be somewhat effective in improving family functioning and maternal well-being in deprived areas. SSLPs appear to have generated some beneficial effects for family functioning and maternal well-being, but not upon child outcomes by the time children are 7 years old. It may be that the parenting effects discerned by age 3 are too little, too late, to produce the improvements necessary for language development that could lead to subsequent educational and social improvement.

It is clear that further improvements in services are desirable, and a further lesson to be learnt from the evaluation is that when developing community-based programmes it is important to ensure that there is the correct balance between professional and community involvement and that the best use is made of available evidence and expertise. In particular, language development in the early years underpins both cognitive and social development. Hence if children's centres are to have an observable impact upon school readiness greater emphasis needs to be given to improving children's language development. It is likely that the failure of SSLPs to produce any improvement in child language development underpins the later absence of SSLP effects upon child outcomes.

It will be some time before the longer term goals of the programme (to improve life course trajectories) can be realised, and hence the final verdict on Sure Start Local Programmes awaits further evaluation. It is possible these beneficial effects on parenting may produce beneficial effects upon child outcomes when the children are older. Other interventions that have affected parenting report long-term beneficial effects upon adolescent criminality and substance abuse (Olds et al., 1999). It may be that the parenting effects discerned by age 3 are too little too late to produce the improvements necessary for language development that could affect subsequent educational and social improvement. Other NESS work has indicated how high quality early childhood education and care that can be provided by children's centres might improve child development outcomes (Melhuish et al., 2010a). Also where children have more effective home learning environments and better parenting, which are improved by SSLPs, these can improve child development outcomes in the long term. It would appear that Sure Start Children's Centres are well-placed to provide improved integrated services to help support the most disadvantaged children and families and potentially assist in narrowing the gap between the disadvantaged and the more advantaged. However, they will need to focus more directly on improvements to young children's daily experience, which is a primary engine of child development, if they are to improve child outcomes.

The current Government[6] has re-iterated that the core purpose of Sure Start Children's Centres is to improve outcomes for young children and their families, with a particular focus on the most disadvantaged, so children are equipped for life and ready for school, no matter what their background or family circumstances. In pursuing this purpose children's centres are continuing to evolve, and the Government has set out its vision for the services that should be on offer for parents, children and families in the early years in the document *Supporting Families in the Foundation Years*[7], which describes the system needed to make the Government's vision a reality and explains the role of commissioners, leaders and practitioners across the range of services for families in these years. In considering how these

[6] DfE 26th April 2012: http://www.education.gov.uk/childrenandyoungpeople/earlylearningandchildcare/a00191780/core-purpose-of-sure-start-childrens-centres

[7] http://www.education.gov.uk/childrenandyoungpeople/earlylearningandchildcare/early/a00192398/supporting-families-in-the-foundation-years

policies are implemented for children's centres the project *Evaluation of Children's Centres in England*[8] is underway and this evaluation will report on the ongoing impact of children's centres.

[8] See http://www.education.ox.ac.uk/research/fell/research/evaluation-of-children-centres-in-england-ecce/ and http://www.natcen.ac.uk/study/evaluation-of-children's-centres-in-england)

REFERENCES

Barnes, J., (2007) Targeting Deprived Areas: The Nature of the Sure Start Local Programme Neighbourhoods In J. Belsky, J. Barnes, & E. Melhuish, (Eds.). *The National Evaluation of Sure Start: Does Area-based Early Intervention Work?* Bristol, UK: The Policy Press.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, *333*, 975-78.

Barrett, H. (2010). *The delivery of parent skills training programmes: Meta-analytic studies and systematic reviews of what works best*. London: Family & Parenting Institute. http://www.familyandparenting.org/Resources/FPI/Documents/TheDeliveryOfParentSkillsTrainingProgrammes.pdf.

Belsky, J., Melhuish, E., Barnes, J., Leyland, A.H., Romaniuk, H., & the NESS Research Team. (2006). Effects of Sure Start Local Programmes on Children and Families: Early Findings from a Quasi-experimental, Cross-sectional Study. *BMJ, 332*, 1476-1578 (BMJ online: http://bmj.com/cgi/doi/10.1136/bmj.38853.451748.2F).

Campbell, F.A., & Ramey, C.T. (1994). Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families. *Child Development, 65,* 684-689.

DfE, (2011) Statistical First Release: Provision for Children Under Five Years of Age in England - June 2011: http://www.education.gov.uk/rsgateway/DB/SFR/s001009/index.shtml

Frost, M., & Harper, G. (2007). The challenge of profiling communities. In J. Belsky, J. Barnes, & E. Melhuish, E. (Eds.). *The National Evaluation of Sure Start: Does Area-based Early Intervention Work?* Bristol, UK: The Policy Press.

Hill, J.L., Waldfogel, J., Brooks-Gunn, J. & Han, W. (2005). Maternal employment and child development: A fresh look using new methods. *Developmental Psychology, 41*, 833-850.

HM Treasury (1998a). *Modern Public Services for Britain: Investing in Reform. Comprehensive Spending Review: New Public Spending Plans.* London: HMSO.

HM Treasury (1998b). *Comprehensive Spending Review: Cross Departmental Review of Provision for Young Children.* London: HMSO.

Love, J., Kisker, E. E., Ross, C. M. Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni, A. S., Brady-Smith, C. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start. Volume 1: Final Technical Report*. Princeton, NJ: Mathematica Policy Research Inc. Available at: http://www.mathematica-mpr.com/PDFs/ehsfinalvol1.pdf

Magnuson, K.A., Ruhm, C., & Waldfogel, J. (2007) Does Pre-kindergarten Improve School Preparation and Performance? *Economics of Education Review, 26*, 33–51.

McCartney, K., Bub, K., & Burchinal, M. (2006). Selection, detection and reflection. In K. McCartney, Burchinal, M.R., & Bub, K. (Eds.), Best practices in quantitative methods (pp. 105-126). *Monographs of the Society for Research in Child Development, Serial 285, Vol. 71*.

Melhuish, E.C., Lloyd, E., Martin, S. and Mooney, A. (1990) Type of childcare at 18 months: II Relations with cognitive and language development. *Journal of Child Psychology and Psychiatry, 31,* 861-870.

Melhuish E. & Hall D. (2007). The policy background to Sure Start. In: Belsky J, Barnes J, Melhuish E, Eds. *The National Evaluation of Sure Start: Does Area-based early intervention work?* Bristol: Policy Press, pp 3-21.

Melhuish, E. Belsky, J., Leyland, A.H., Barnes, J. & NESS Research Team (2008a). Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study. *Lancet, 372*, 1641-1647.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. & Malin, A. (2008b). Preschool influences on mathematics achievement. *Science, 321,* 1161-1162.

Melhuish, E.C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., & Phan, M. (2008c) Effects of the Home Learning Environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64,* 95-114**.**

Melhuish, E., Belsky, J., MacPherson, K., & Cullis, A. (2010a). *The quality of group childcare settings used by 3-4 year old children in Sure Start local programme areas and the relationship with child outcomes*: Research report DFE-RR068. London: DfE. ISBN 978-1-84775-828-6. http://publications.education.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DFE-RR068&

Melhuish, E., Quinn, L., Sylva, K., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (2010b). *Pre-school Experience and Key Stage 2 performance in English and Mathematics.* Belfast: Dept for Education, Northern Ireland. Available at:

Melhuish, E. (2010). *Why children, parents and home learning are important*. In K. Sylva, E. Melhuish, P. Sammons, I. Siraj-Blatchford, B. Taggart, (Eds.) (2010). Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education Project. London: Routledge.

Melhuish E. C. (2011). Preschool matters*. Science, 333,* 299-300.

NESS Research Team (2005). *Early impacts of Sure Start Local Programmes on Children and Families*. Report 13. London: Department for Education and Skills,

NESS Research team (2008). *The impact of Sure Start Local Programmes on three year olds and their families.* London: DCSF.

NESS (2010). *The impact of Sure Start Local Programmes on five year olds and their families*: Research report DFE-RR067. London: DfE. ISBN 978-1-84775-826-2. http://publications.education.gov.uk/eOrderingDownload/DFE-RR067.pdf

ODPM. (2004), *The English Indices of Deprivation 2004 (revised)*. London: ODPM.

Olds, D. L., Henderson, C.R., Kitzman, H., Eckenrode, J. J., Cole, R. E., & Tatelbaum, R. C. (1999). "Prenatal and Infancy Home visitation by Nurses: recent Findings," *Future of Children, 9*, 44-66.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, *2nd Ed.* Cambridge: CUP.

Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent Effects of Early Childhood Education on High-risk Children and their Mothers. *Applied Developmental Science, 4*(1), 2-14

Reynolds, A.R., Temple, J.A., Ou, S., Arteaga, I.A., & White, B.A.B. (2011). School-Based Early Childhood Education and Age-28 Well-Being: Effects by Timing, Dosage, and Subgroups. *Science, 333*, 360-364.

Rosenbaum, P. R., Rubin, D. B. (1983). "The central role of the propensity score in observational studies for causal effects". *Biometrika, 70* (1): 41–55.

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys.* London: Wiley.

Rubin, D.B. (1997). Estimating causal effects from large datasets using propensity scores. *Annals of Internal Medicine, 127*, 757-763.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., and Hunt, S. (2008). *Influences on Children's Attainment and Progress in Key Stage 2: Cognitive outcomes in Year 6.* London: DCSF.

Schrader-McMillan, A., Barnes, J. & Barlow, J. (2012, in press) *Primary study evidence on effectiveness of interventions (home, early education, child care) in promoting social and emotional wellbeing of vulnerable children under 5*. London: NICE.

Schweinhart, L.J., Barnes, H.V., Weikart, D.P., Barnett, W.S., & Epstein, A.S. (1993). *Significant Benefits: The High/Scope Perry Preschool Study through Age 27.* Monographs of the High/Scope Educational Research Foundation, Number 10, Ypsilanti, Michigan: High/Scope Press.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B., (Eds.) (2010). *Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education Project*. London: Routledge.

Tunstill, J., Meadows, P., Akhurst, S., Allnock, D., Chrysanthou, J, Garbers, C. & Morley, A (2005). Implementing Sure Start Local Programmes: An overview of the first four years. London: DfES.

Vandell, D.L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD Early Child Care Research Network (2010). The Do Effects of Early Child Care Extend to Age 15 Years? Results From the NICHD Study of Early Child Care and Youth Development. *Child Development, 81*, 737-756.

APPENDICES

Appendix A: Procedures for Propensity Matching

In order to be able to implement propensity scoring analysis, it was essential to determine how many of the 138 non-SSLP (MCS) areas were sufficiently comparable to the SSLP areas to be useful in an analysis. Toward this end, the 138 identified MCS areas were compared with the 150 SSLP areas on 85 indices of deprivation and other socio-demographic variables obtained from administrative sources (see Table A.1). These variables were used in an attempt to distinguish between SSLP and MCS areas. The table below displays the mean and standard deviation for each of the 85 area deprivation variables in the 150 SSLP and 138 MCS areas. Also shown is the standardised percentage difference – the difference between the two means expressed as a percentage of the average standard deviation.

(The average standard deviation is $\sqrt{½\left(s_{SS}^2 + s_{MCS}^2\right)}$.)

Finally, Table A.1 shows a two sample t-statistic testing for the significance of a difference between SSLP and MCS areas and its p-value.

Table A.1: Mean and Standard Deviations of SSLP and MCS Areas on 85 Area Deprivation Variables

N	Deprivation Indicator	SSLP mean N=150	SSLP SD	MCS mean N=133	MCS SD	Stand % diff	t-statistic	p-value
1	**% lone parent families**	27.10	7.98	19.35	7.49	100.15	8.48	0.000
2	**% inflow of all households with children**	7.25	1.84	6.59	2.44	30.39	2.59	0.010
3	**% outflow of all households with children**	7.78	1.85	6.57	2.31	57.55	4.90	0.000
4	% Europe	91.08	11.18	92.03	11.74	-8.24	-0.70	0.485
5	**% Asian Bangladeshi**	1.73	6.14	0.49	1.85	27.38	2.28	0.023
6	**% Asian Indian**	1.79	3.33	3.97	9.77	-29.99	-2.59	0.010
7	**% Asian Pakistani**	4.66	12.13	1.91	5.46	29.18	2.44	0.015
8	% Black African	2.66	5.65	1.56	3.73	23.03	1.94	0.054
9	% Black Caribbean	2.18	3.83	1.74	3.59	11.81	1.00	0.318
10	**% Chinese**	0.48	0.64	0.33	0.39	27.14	2.28	0.023
11	**% mixed**	1.83	1.71	1.46	1.29	24.17	2.04	0.043
12	% other	1.12	1.68	1.29	2.60	-7.42	-0.63	0.526
13	% white British	81.27	22.90	85.00	21.35	-16.83	-1.43	0.155
14	% white other	2.29	3.25	2.21	2.47	2.80	0.24	0.814
15	**% of all people LLTI**	21.41	3.89	17.98	3.95	87.44	7.42	0.000
16	**% of people working or seeking with LLTI**	8.74	1.11	7.63	1.30	91.85	7.81	0.000
17	**% no working parents with children**	29.38	10.01	15.92	8.35	146.10	12.34	0.000
18	**% unemployed**	6.08	1.87	3.82	1.75	124.70	10.56	0.000
19	% economically active ft student	2.48	1.36	2.32	0.97	13.37	1.13	0.261
20	**% long term unemployed**	3.76	1.32	2.20	1.19	124.12	10.50	0.000
21	**% all managerial**	16.26	6.43	23.09	8.01	-94.03	-8.01	0.000
22	**% lower managerial**	12.08	4.16	16.64	4.64	-103.57	-8.80	0.000
23	**% intermediate**	7.40	1.88	9.30	2.55	-85.00	-7.25	0.000
24	**% small employers**	4.97	1.82	6.90	3.07	-76.22	-6.53	0.000
25	% lower supervisory and technical	7.82	1.94	8.24	2.26	-20.06	-1.71	0.089
26	**% all routine**	27.86	6.09	24.95	6.92	44.54	3.79	0.000
27	**% never worked and long term unemployed**	8.49	5.02	4.18	2.85	105.79	8.87	0.000
28	**% not classified**	27.20	4.72	23.31	4.75	82.15	6.97	0.000

29	**% vacant household spaces**	4.71	3.38	2.63	2.11	73.99	6.22	0.000
30	Of all occupied hhold spaces: % unshared	99.63	0.53	99.75	0.57	-21.29	-1.81	0.072
31	**% of all households owned**	47.81	14.57	63.30	16.26	-100.33	-8.53	0.000
32	**% all households social and council rented**	39.38	15.48	25.39	16.11	88.51	7.51	0.000
33	**% over 1.5 persons per room**	1.24	1.68	0.82	1.28	28.65	2.42	0.016
34	**% of all hholds with no dependent children**	66.41	5.91	67.96	6.95	-24.14	-2.05	0.041
35	% Christian	65.54	15.82	68.58	13.66	-20.59	-1.74	0.083
36	% Buddhist	0.29	0.36	0.28	0.39	2.44	0.21	0.836
37	**% Hindu**	0.73	1.51	2.68	7.19	-37.36	-3.23	0.001
38	% Jewish	0.30	1.77	0.28	0.83	1.75	0.15	0.884
39	**% Muslim**	8.36	14.91	3.56	7.42	40.77	3.41	0.001
40	% Sikh	0.56	1.50	1.18	4.45	-18.65	-1.61	0.109
41	% any other religion	0.27	0.20	0.34	0.42	-20.29	-1.74	0.082
42	% no religion	14.97	5.97	15.09	6.18	-2.05	-0.17	0.862
43	**% religion not stated**	8.98	1.78	7.99	1.75	56.22	4.76	0.000
44	**% no qualifications**	41.16	8.43	33.04	8.71	94.85	8.05	0.000
45	**% of under 24 with no qualifications**	24.11	6.84	18.19	6.21	90.62	7.67	0.000
46	**standardised LLTI males (per 100)**	21.82	4.38	15.22	4.12	154.91	13.12	0.000
47	**standardised LLTI females (per 100)**	20.72	3.75	15.53	3.71	139.41	11.82	0.000
48	**% of all people aged 0-4**	7.43	1.49	6.53	1.55	59.08	5.01	0.000
49	**% of all people aged 65+**	13.38	3.21	14.79	4.91	-34.02	-2.91	0.004
50	**% hholds all pensioners**	20.33	4.40	22.05	7.34	-28.33	-2.43	0.016
51	**% people in hholds with no car or van**	36.45	11.58	21.76	10.52	132.80	11.24	0.000
52	**% of aged 16+ ft students at term time**	6.75	4.45	5.48	2.87	34.07	2.86	0.005
53	**% age 15-24 in ft educ living away term**	2.52	1.67	4.77	4.27	-69.50	-5.98	0.000
54	**Weighted pay check mean**	23.26	4.43	28.58	5.22	-109.86	-9.35	0.000
55	**% HH income < 60% national median**	37.57	8.23	27.38	7.75	127.59	10.80	0.000
56	**IMD score 2004**	43.61	12.72	24.80	9.83	165.44	13.95	0.000
57	**IMD crime score 2004**	0.75	0.52	0.16	0.60	105.66	8.99	0.000
58	**IMD education score 2004**	45.13	17.03	27.30	14.12	114.04	9.63	0.000
59	**IMD employment score 2004**	0.20	0.07	0.11	0.06	143.52	12.12	0.000
60	**IMD environment score 2004**	33.02	16.42	23.25	12.22	67.47	5.69	0.000
61	**IMD health score 2004**	1.02	0.55	0.24	0.55	143.32	12.15	0.000
62	IMD housing score 2004	21.10	10.66	22.37	10.93	-11.70	-0.99	0.322
63	**IMD IDAC score 2004**	0.42	0.13	0.24	0.12	143.22	12.12	0.000
64	**IMD IDAOP score 2004**	0.29	0.10	0.18	0.07	120.79	10.18	0.000
65	**IMD income score 2004**	0.30	0.10	0.16	0.08	152.63	12.88	0.000
76	**GO EE**	0.05	0.23	0.13	0.34	-26.84	-2.29	0.023
77	GO EM	0.08	0.27	0.09	0.29	-5.02	-0.43	0.670
78	GO LO	0.16	0.37	0.17	0.38	-3.72	-0.32	0.753
79	**GO NE**	0.13	0.33	0.04	0.19	33.41	2.80	0.005
80	GO NW	0.19	0.40	0.14	0.35	12.89	1.09	0.276
81	**GO SE**	0.07	0.26	0.19	0.39	-34.51	-2.95	0.003
82	GO SW	0.07	0.26	0.08	0.27	-2.39	-0.20	0.839
83	GO WM	0.11	0.32	0.08	0.27	11.36	0.96	0.338
84	GO YH	0.13	0.33	0.07	0.26	18.11	1.53	0.128
85	**Rural**	0.05	0.23	0.14	0.35	-28.90	-2.47	0.014

The fact that so many of the variables differed between SSLP and MCS areas—with significance denoted by bold type, was not important in itself; because as long as there was a reasonable overlap between the two samples, it should prove possible to adjust for the difference. For example, the total IMD score differs between samples with a mean (min-max) of 43.61 (14.74 – 76.13) in SSLP areas and 24.80 (13.79 – 71.81) in the MCS areas. There was considerable overlap, but it needed to be determined whether it was sufficient to adjust for differences in IMD total score between the two area types.

The data displayed in Table A.1 show, not surprisingly, that SSLP populations were, in general, more disadvantaged than the comparison population drawn from deprived MCS areas. The ethnic and religious mix of the areas differ, with SSLP

areas having lower proportions of Asian Indians and Hindus than the MCS areas and higher proportions of Asian Bangladeshi, Asian Pakistani, Chinese, mixed and Muslim populations. The geographical spread shows that a higher proportion of the SSLP areas are in the North East and more of the comparison areas are in the East of England, the South East and in rural areas. For identifiability the categories % white British, % all managerial, and % no religion were excluded when creating the propensity score.

To determine the propensity of an area to be a SSLP impact study area, logistic regression analysis was conducted with the outcome being "*SSLP or comparison area*" based on the variables listed in Table A.1 with the exception of regions and the categories omitted listed in the previous paragraph. The analysis was conducted using standardised scores so that the relative importance of each variable could be evaluated. Increasing the number of variables in the model naturally tended to increase the ability of the model to discriminate correctly between SSLP and comparison areas. The most influential variables were total IMD score (with SSLP areas tending to have higher IMD scores) and the proportion of Asian Indians (with SSLP areas tending to have lower proportions; see Table A.2).

Table A.2: Logistic Regression Results--Percent Correct Classification of SSLP and MCS Areas

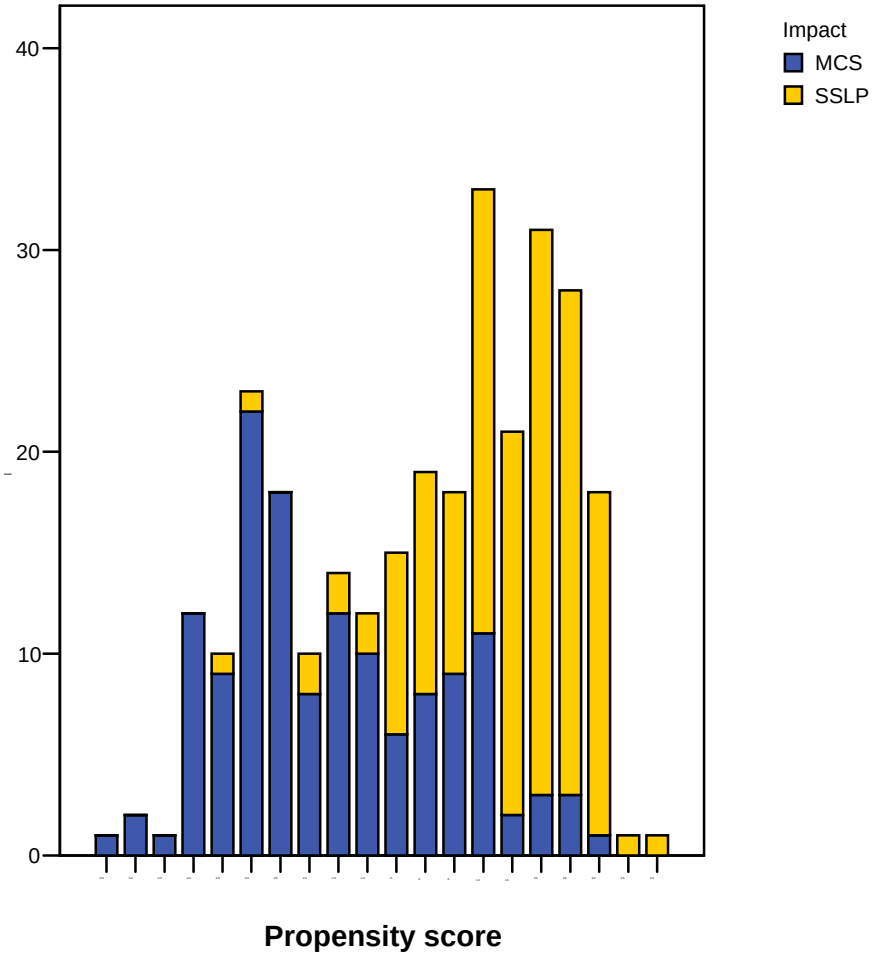
Model	% correct SSLP	% correct MCS	% correct total	Stand coeff
IMD score 2004	82.0	77.5	79.9	2.100
+ % Asian Indian	82.0	79.7	80.9	-0.408
+ % of people working or seeking with LLTI	82.0	77.5	79.9	0.361

In addition to the above terms included in the logistic regression, including a quadratic term produced significant improvement in model fit. This model correctly classified (with a 50% cut-off) 80.2% of the areas as SSLP and MCS areas. The extent of the overlap on the propensity score is displayed in Figure A.1, a stacked-bar chart, which shows that MCS areas have a heavy left-hand tail. This can in part explain the slightly higher standard deviation (SD) in the propensity score for the comparison areas (see Table A.3). Notably, there were few comparison areas with high propensity scores.

Table A.3: Propensity-score Descriptive Statistics for 150 SSLP and 138 MCS Areas

	N	Mean	SD	Min	Max
SSLP	150	1.42	1.31	-3.43	3.81
MCS	138	-1.63	1.89	-5.70	3.17

Figure A.1: Distribution of Propensity Scores as a Function of SSLP (NESS) and non-SSLP (MCS)



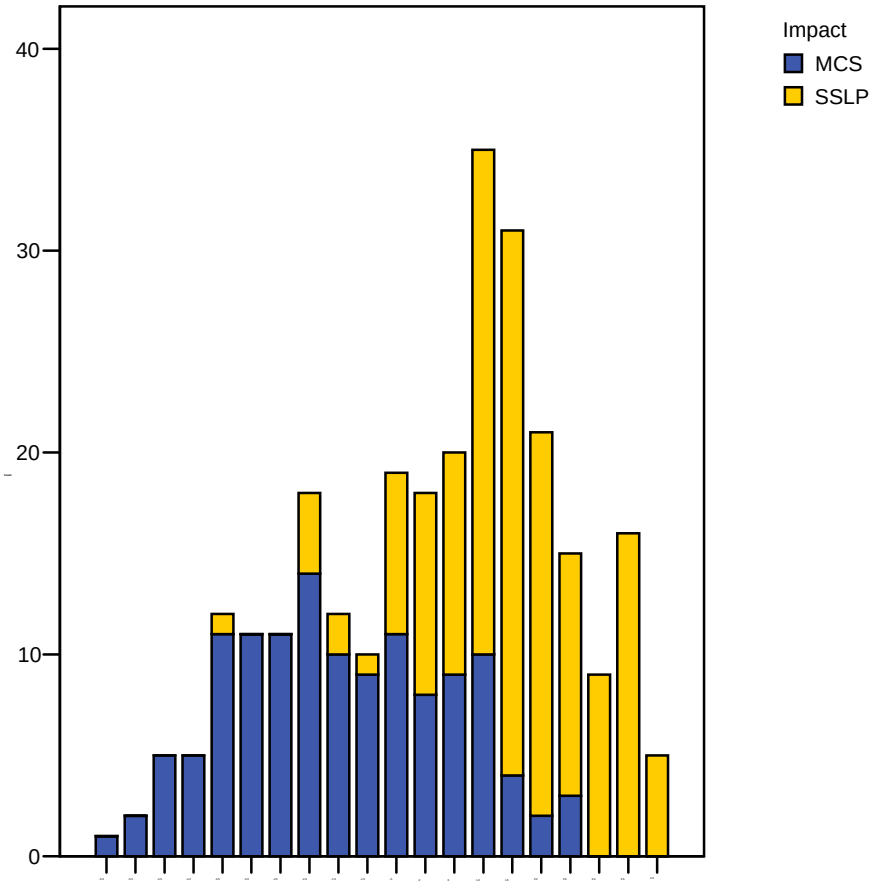
The exclusion of 4 MCS comparison areas with extreme (low) propensity scores yielded 284 areas that showed some overlap between the SSLP and MCS areas (with propensity scores ranging from -4.22 to 3.17). The difference between the means of these two groups remained still significant (see Table A.4).

Table A.4: Propensity-score Descriptive Statistics for 150 SSLP and 134 MCS Areas

	N	Mean	SD	Min	Max
SSLP	150	1.42	1.31	-3.43	3.81
MCS	134	-1.52	1.82	-4.22	3.17

With these four outlying MCS areas excluded the propensity score was re-created. Repeating the process on the subset of areas resulted in a different propensity score – both in terms of the regression coefficients and the variables used to distinguish between the two sets of areas. Cycling through this process resulted in the exclusion of another eight MCS areas, leaving a total of 276 areas (150 SSLP, 126 MCS). Despite clear differences in the distributions, there was the (necessary) overlap on the propensity score (see Figure A.2).

Figure A.2: Revised Distribution of Propensity Scores as a Function of SSLP (NESS) and non-SSLP (MCS)



The distribution of propensity scores across the two studies, as displayed in Figure A.2, indicated that exact matching could not be achieved. Any form of one-to-one propensity-score matching – such as nearest neighbour – would require the exclusion of some cases, given that there were more SSLP than MCS areas, and could therefore lead to the biases associated with incomplete matching. Some of the "nearest neighbours" would also still be fairly dissimilar. The alternative propensity-matching strategy in this situation was followed, namely, dividing the distribution along the propensity score into strata. Monte Carlo simulations suggest that a weighted analysis using five strata with adjustment for the propensity score within each stratum should result in little bias in estimating effects of SSLPs.

Thus, using five equal strata we proceeded to look for significant main effects of SSLPs or interactions with the strata for each of the covariates. A propensity score based on two variables together with one quadratic term correctly classified 79.1% of the areas as SSLP or MCS. It turned out, however, that when stratified on the basis of this propensity score the IMD IDAOP score 2004 was unbalanced within the strata. A two-way analysis of variance showed that the marginal mean of this variable was higher within SSLP than MCS areas (see Table A.5).

Table A.5: Marginal Means of IMD IDAOP Score for Propensity-Score Strata for SSLP and non-SSLP Areas

	SSLP		Non-SSLP		
	Mean	SE	Mean	SE	p
Marginal mean	0.245	0.010	0.212	0.009	0.014
Strata:					
1	0.165	0.046	0.134	0.009	0.006
2	0.194	0.016	0.176	0.011	
3	0.228	0.011	0.246	0.014	
4	0.261	0.010	0.253	0.022	
5	0.378	0.009	0.250	0.033	

Following the inclusion of this variable (along with others) in the propensity score, the differences between SSLP and non-SSLP areas were no longer significant, thus yielding balance within all strata for all variables. That is, two-way analyses of variance indicated that there were no significant main effects or interactions. The results for the IMD IDAOP score 2004 are displayed in Table A.6.

Table A.6: Revised Means of IMD IDAOP Score for Propensity-Score Strata for SSLP and non-SSLP Areas

	Sure Start		MCS		
	Mean	SE	Mean	SE	p
Marginal mean	0.243	0.010	0.214	0.014	0.095
Strata:					
1	0.165	0.046	0.135	0.009	0.170
2	0.197	0.017	0.178	0.010	
3	0.218	0.011	0.240	0.014	
4	0.260	0.010	0.265	0.021	
5	0.376	0.009	0.250	0.065	

The final propensity score (based on standardised variables) is displayed in Table A.7 and the distributions of SSLP and non-SSLP areas across the five propensity-score strata are shown in Table A.8.

Table A.7: Final Propensity Score

Variable	Standard coefficient
Constant	0.834
% Intermediate workers	0.055
Standardised LLTI (males)	0.793
IMD score 2004	1.313
IMD employment score 2004	0.502
IMD IDAOP score 2004	0.277
(Standardised LLTI (males))^2	-0.066
(IMD employment score 2004)^2	-1.044
(IMD employment score 2004)*(IMD IDAOP score 2004)	0.932

Table A.8: Distributions of SSLP and non-SSLP Areas for Five Propensity Strata, including Sample Sizes

Propensity Stratum	SSLP N Areas	non-SSLP N Areas	N children
1	2	53	1,041
2	15	40	970
3	33	22	818
4	45	10	565
5	55	1	21
Total	150	126	3,415

Importantly, the different distribution of non-SSLP and SSLP areas in each stratum can be accounted for by weighting the analysis. Thus, each non-SSLP area is weighted relative to the ratio of SSLP to non-SSLP areas within that stratum, which is equivalent to weighting by the selection probability. Whereas SSLP areas in the sample have a weight of 1, the weight attached to each non-SSLP area in the 1st stratum would be 0.038 (i.e., 2/53); if this stratum were to be included in the analysis; this would ensure that less weight would be given to the (many) non-SSLP areas with the lowest propensity scores and that are least typical of SSLP areas. In the 5th stratum the weight attached to the non-SSLP area would be 55 (i.e., 55/1), the increase in weight compensating for the fact that only one non-SSLP area is in this stratum.

The strata are based on a scale (the propensity score) that rates areas on their tendency to be SSLP areas; using the variables listed above it is clear that there were areas at both extremes that are clearly distinguished. That is, only one non-SSLP area had the characteristics “most like SSLP areas” and only two SSLP areas had those “least like SSLP areas.” It is important to note that even though the “most like SSLP” areas tend to be more deprived, the propensity score is not a straightforward measure of deprivation. In fact, had we, for example, used the IMD score to stratify the data, we would have found increased numbers to make comparisons among, for example, the most deprived areas (see Table A.9).

Table A.9: Distribution of SSLP and non-SSLP Areas Using IMD Data to Stratify Areas

IMD score	Sure Start	Non-SSLP	
Stratum	N Areas	N Areas	N children
1	3	52	1,010
2	17	38	821
3	33	22	1,098
4	45	10	430
5	52	4	56
Total	150	126	3,415

The end result is that we succeeded in identifying in the MCS a sample of 3,415 children nested in 126 areas that can be used (stratified) as a non-SSLP comparison group (see Table A.8). Of these MCS children 2,799 were seen at both 9 months and 3 years of age, and of these 2537 children had provided sufficient data to be used in analyses. When stratified there proved to be a good balance within each stratum for all the area variables. The analysis is weighted to take into account the unequal distribution of SSLP and non-SSLP areas across the different strata. It must be noted that it would prove difficult in the final analyses to examine SSLP effects among the stratum characterised by the greatest propensity to be SSLP areas and which will include a lot of the most deprived areas because, in this stratum, the non-SSLP only provides 21 children in one comparable area.

Appendix B: Comparison of Children/Families Seen and Not Seen at 7 years.

In the following tables of Appendix B the distribution of characteristics for the sample seen at 7 years is given in the first (left) part of the table; the characteristics of the sample not seen at 7 years (attrition) is given in the middle column; and the results of tests of the differences between those seen and not seen at 7 years are given in the right part of the table. Data are broken down by strata 1 to 5. Also the data are presented separately for the Sure Start (NESS) (B1) and non- Sure Start (B2) samples used in the study. Note that some comparisons involve cells with low numbers, e.g. often applies to stratum 1 for NESS sample.

Appendix B: Comparison of Children/Families Seen and Not Seen at 7 years.

Table B.1: Sure Start (NESS) sample - Comparison of Children/Families Seen and Not Seen at 7 years

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Child's Gender											
Male	46.1	47.2	48.5	50.8	52.5	51.4	.24	.47	.22	.38	.34
Female	53.9	52.8	51.5	49.2	47.5	48.6					
Child's Ethnicity											
White	95.4	85.7	85.1	81.6	61.9	77.5	.28	.86	.56	.63	.00
Mixed	1.3	4.2	4.4	4.2	6.2	4.9					
Indian	1.3	1.8	1.2	1.2	1.8	1.0					
Pakistani	-	1.0	2.6	5.3	12.7	6.5					
Bangladeshi	-	0.7	1.1	0.9	6.4	2.2					
Black Caribbean	-	1.8	1.6	1.2	2.2	1.1					
Black Other	-	2.7	1.9	3.7	4.0	4.1					
Other	-	2.2	2.1	1.8	4.8	2.7					
Language in Home											
English Home Language	97.4	90.2	90.0	85.7	70.5	82.8	.97	.14	.79	.06	.00
Other Languages	2.6	9.8	10.0	14.3	29.5	17.1					
Missing	-	-	-	-	-	0.1					
Maternal Age at Birth of Child											
Not teenage	96.1	94.0	92.1	90.4	88.8	85.2	.69	.02	.00	.00	.00
Teenage (< 20 years)	3.9	6.0	7.9	9.6	11.2	14.7					
Missing	-	-	0.1	0.1	0.1	0.1					

Table B.1 (continued): Sure Start (NESS) sample

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Maternal Cognitive Difficulties											
No Difficulties Reported	92.1	91.8	90.9	92.6	88.0	88.1	.25	.31	.11	.01	.08
Has Some Difficulties	7.9	8.2	9.1	7.4	12.0	11.9					
Missing	-	-	0.1	-	0.1	0.1					
Household Deprivation											
Above poverty line[+]	82.4	63.9	55.7	53.9	34.8	40.1	.07	.32	.01	.00	.00
Below poverty line[+]	17.6	36.1	44.3	46.1	65.2	56.4					
Missing	2.6	2.5	3.2	4.5	4.5	3.5					
Highest Occupation in Household											
Management/Professional	45.2	30.3	24.6	22.3	14.1	15.6	.33	.02	.00	.00	.00
Intermediate	15.1	9.2	8.5	10.0	6.9	6.9					
Small Employer	4.1	10.8	6.3	6.0	5.2	5.9					
Lower Supervisory/Technical	11.0	9.2	10.6	8.8	6.4	7.0					
Semi-Routine	9.6	9.2	11.8	13.1	15.1	12.5					
Routine	4.1	8.3	10.1	8.8	8.9	8.4					
Unemployed	11.0	23.0	28.1	31.1	43.4	42.5					
Missing	3.9	-	1.4	1.1	1.4	1.3					

[+]Poverty line is £210 per week, taken from the financial year 2004-2005

Table B1 (continued): Sure Start (NESS) sample

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Lone Parent											
Not Lone Parent	86.8	77.8	76.9	71.5	64.5	65.5	.92	.56	.00	.00	.01
Lone Parent	13.2	22.2	23.1	28.5	35.5	34.4					
Missing	0.1	-	-	-	-	0.1					
Work Status Household											
Working Household	89.5	77.0	72.3	69.3	55.7	57.5	.48*	.12	.00	.00	.00
Workless Household	10.5	23.0	27.7	30.7	44.2	42.5					
Missing	-	-	0.1	-	0.1	0.1					
Highest Education in Household											
Degrees/Higher Education	59.2	33.5	29.3	29.2	20.1	20.2	.09	.00	.00	.00	.00
A level	17.1	27.3	27.2	26.8	27.5	24.9					
O level / GCSE	14.5	22.8	24.1	21.7	21.8	23.0					
Other	1.3	7.3	7.0	6.5	6.7	7.2					
None	7.9	9.0	12.4	15.8	23.9	24.5					
Missing	-	-	0.3	-	0.1	0.3					

Table B.2: Non Sure Start sample - Comparison of Children/Families Seen and Not Seen at 7 years

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Child's Gender											
Male	50.5	48.7	50.3	48.8	-	52.7	.19	.71	.73	.14	-
Female	49.5	51.3	49.7	51.2		47.3					
Child's Ethnicity											
White	91.8	81.2	71.5	71.0	-	73.7	.06	.20	.19	.76	-
Mixed	2.3	5.2	4.2	4.6	-	5.9					
Indian	0.7	3.6	1.9	12.5	-	4.9					
Pakistani	1.5	4.2	11.0	2.7	-	4.7					
Bangladeshi	1.0	0.5	5.7	0.3	-	2.2					
Black Caribbean	1.1	0.8	1.7	2.7	-	2.2					
Black Other	0.5	1.7	1.9	3.0	-	3.2					
Other	1.1	2.8	1.9	3.0	-	3.2					
Language in Home											
English Home Language	94.7	87.3	78.6	78.4	-	81.0	.06	.54	.49	.04	-
Other Languages	5.3	12.7	21.4	21.6	-	19.0					
Maternal Age at Birth of Child											
Not teenage	95.0	93.6	89.6	92.1	-	88.0	.04	.05	.22	.08	-
Teenage (< 20 years)	5.0	6.4	10.4	7.9	-	11.8					
Missing	-	-	-	-		0.2					

Table B.2 (continued): Non Sure Start sample

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Maternal Cognitive Difficulties											
No Difficulties Reported	91.4	89.9	90.2	94.2	-	90.6	.30	.54	.53	.49	-
Has Some Difficulties	8.6	10.1	9.8	5.8	-	9.1					
Missing	0.1	0.3	0.0	0.3	-	0.3					
Household Deprivation											
Above poverty line[+]	81.8	68.8	59.1	51.2	-	46.5	.00	.00	.00	.15	-
Below poverty line[+]	18.2	31.2	40.9	48.8	-	41.9					
Missing	7.3	8.3	10.2	8.8	-	11.6					
Highest Occupation in Household											
Management/Professional	49.9	35.1	27.1	24.1	-	18.7	.00	.00	.01	.10	-
Intermediate	11.1	8.4	8.5	8.0	-	7.7					
Small Employer	8.9	10.2	10.7	6.6	-	5.7					
Lower Supervisory/Technical	9.8	11.9	10.4	10.1	-	8.2					
Semi-Routine	6.4	9.6	13.6	12.9	-	8.4					
Routine	5.9	8.1	6.8	12.2	-	6.7					
Unemployed	8.1	16.7	23.0	25.9	-	26.4					
Missing	8.0	8.8	12.3	12.8	-	18.0					

[+]Poverty line is £210 per week, taken from the financial year 2004-2005

Table B.2 (continued): Non Sure Start sample

Characteristic	Percentage in Each Stratum						Comparison between Each Stratum and Attrition				
	Strata: SSLP Community Like					Attrition	p-values				
	least 1	2	3	4	most 5		least 1	2	3	4	most 5
Lone Parent											
Not Lone Parent	91.3	87.1	85.6	81.7	-	76.4	.05	.00	.00	.03	-
Lone Parent	8.7	12.9	14.4	18.3	-	23.6					
Work Status Household											
Working Household	92.2	84.0	78.5	75.6	-	70.5	.00	.00	.00	.09	-
Workless Household	7.8	16.0	21.5	24.4	-	29.5					
Missing	4.1	4.6	6.4	7.6	-	10.4					
Highest Education in Household											
Degrees/Higher Education	45.5	36.3	25.5	26.6	-	23.6	.09	.01	.52	.04	-
A level	34.1	34.8	38.7	33.2	-	26.3					
O level / GCSE	15.7	18.8	21.2	26.2	-	19.7					
Other	0.4	3.3	3.6	4.2	-	2.4					
None	4.3	6.8	11.1	9.8	-	11.3					
Missing	7.3	8.3	11.7	12.8	-	16.8					

Appendix C: Description of Outcome Variables

Physical Health	
BMI	Body Mass Index: weight in kgs. divided by square of height in metres, and then standardised by age and gender
General health	Respondent's rating of the child's general health
Child Educational Development - All teacher ratings – see also section 2.6	
	Every child in a state school is assessed at Key Stage 1 (KS1). This occurs towards the end of the school year when the child becomes 7 years old. Typically this will be the child's third year in primary school. The KS1 assessments used in analysis are reading, writing, mathematics and science.
KS1 reading	• KS1 Points score standardised by year of assessment
KS1 writing	• KS1 Points score standardised by year of assessment
KS1 mathematics	• KS1 Points score standardised by year of assessment
KS1 science (overall)	• KS1 Points score standardised by year of assessment
Child Social and Emotional Development	
Emotional dysregulation	A construct of items related to: temper tantrums, fighting, bullying, lies, cheating, restlessness, distractibility, mood swings, overexcitement, frustration.
Positive social behaviour	A construct of items related to: having friends, being liked, considerate, sharing, helpful, kind, plays easily with others, cooperative.
Internalisation	A construct of items related to: often has headaches, worried, unhappy, nervous, fearful, solitary, picked on, gets on better with adults than children.
Self-regulation	A construct of items related to: works things out for self, does not need much help, sees things through, chooses activities on their own, persists even when something is difficult, and can move to a new game after playing with a toy or game.
Parenting and Family Functioning	
Harsh Discipline	A construct of: frequency of (reported) ignoring child, smacking, shouts, sends to bedroom/naughty chair, takes treats away, tells off child
Home chaos	Mother reports of extent that household is disorganized.
Home Learning Environment	A construct of 8 items for the frequency of learning opportunities provided to child in home; taken to library, helped with reading, help with writing, help numbers/maths, reads to child, tells stories, musical activities/songs/rhymes with child, child paints and draws

Maternal Well-Being	
Life Satisfaction	How satisfied/dissatisfied respondent is about the way life has turned out.
Self-rated depression	A construct of frequency of:depressed so nothing can cheer you up; feel hopeless; restless or fidgety; everything an effort; feel worthless; feel nervous

Appendix D: Imputation procedure

There is a strategy to overcome the problem that data may be missing in non-random ways and hence bias results. This involves the "imputation" of missing data. Imputation is based on the fundamental premise that tolerably accurate estimates of what a missing value would have been had the information been supplied can be determined using all the data that has been collected. Taking an over-simplified example, knowing a person's age, education level, gender, work status and occupation enables a reasonably accurate prediction of salary, should salary data be missing, using data on all these variables obtained from respondents who also provided salary information. In the current evaluation, statistically sophisticated and widely used multiple-imputation techniques were employed to overcome the possibility of bias in results caused by non-random missing data. This takes into account that, in the above example, we can predict not just one value for the missing salary but a range of plausible values.

Two approaches have been taken for dealing with missing data: case deletion and imputation. Case deletion involves deleting for each outcome measure any individual who has missing data either for the outcome measure or for the demographic or family background characteristics, leaving only cases with complete data. Analysis of data with only complete cases has the drawback that, where we are considering a number of explanatory factors, we may discard quite a large part of the data. Not only is this inefficient, it may result in a subset of data that is small and, if data are missing in a non-random way (i.e. if certain subgroups are more likely to refuse to answer or skip over certain questions), may be unrepresentative of the population as a whole. Imputation of data for a respondent involves filling in the missing values with plausible values based on the known characteristics of that respondent together with the relationship between characteristics observed in the rest of the sample.

Multiple imputation (Rubin, 1987; Schafer, 1997) was used to estimate missing data values using the statistical package IVEware (http://www.isr.umich.edu/src/smp/ive/). Multiple imputation is the process of generating several data sets, analysing these and combining the results. This ensures that we have sufficient variability between imputed values to be able to draw correct inferences. The missing demographic and outcome data were imputed simultaneously for both waves of data. In each strata, the missing data were imputed separately for each study group. Fixed area effects were included to take account of the hierarchical structure of the data. For all the demographic and outcome variables included in the analysis of strata 2-4, 3.6% of the data were missing in the dataset. Table D.1 show the percent missing for each of the demographic and family background characteristic variables by strata and study, for those strata used in the analyses, and also for strata 2-4 by study. Table D.2 shows for each outcome measure the percent of cases with missing outcome measures and incomplete data (missing outcome and/or missing demographic or family background characteristics). For the complete-cases data analysis, the incomplete-data cases are excluded. For the 9-month data, between 8% and 14% of the cases are excluded from any one analysis. For the 3-year data, higher rates of missing data were observed, between 10% and 29% of the cases.

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys.* London: Wiley.
Schafer, J.L. (1997) *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall.

Table D.1: Summary of Demographic Characteristics: Dataset Age 7 yrs for imputation

Characteristic	Sure Start: 7 years							Non Sure Start: 7 years						
	% missing by Strata (N=)					Missing for Strata 2-4 (N=)		% missing by Strata (N=)					Missing for Strata 2-4	
	Strata: SSLP Community Like							Strata: SSLP Community Like						
7 years	Least 1	2	3	4	most 5	N	%	least 1	2	3	4	most 5	N	%
Child's Gender	0.0	0.0	0.0	0.0	0.0	0	0.0	-	0.0	0.0	0.0	-	0	0.0
Child's Ethnicity	0.0	0.0	0.0	0.0	0.0	0	0.0	-	0.0	0.0	0.0	-	0	0.0
Language in Home	0.0	0.0	0.0	0.0	0.0	0	0.0	-	0.0	0.0	0.0	-	0	0.0
Maternal Age at Birth of Child	0.0	0.8	0.7	0.9	1.4	29	0.8	-	1.6	1.5	0.9	-	20	1.4
Maternal Cognitive Difficulties	0.0	0.0	0.1	0.0	0.1	1	0.0	-	0.3	0.0	0.3	-	3	0.2
Household Deprivation	2.6	2.5	3.2	4.5	4.5	130	3.7	-	8.3	10.2	8.8	-	130	9.1
Highest Occupation in Household	3.9	0.0	1.4	1.1	1.4	36	1.0	-	8.8	12.3	12.8	-	156	10.9
Lone Parent	0.0	0.0	0.0	0.0	0.0	0	0.0	-	0.0	0.0	0.0	-	0	0.0
Work Status Household	0.0	0.0	0.1	0.1	0.0	1	0.0	-	4.6	6.4	7.6	-	84	5.8
Highest Education in Household	0.0	0.0	0.3	0.0	0.1	4	0.1	-	8.3	11.7	12.8	-	150	10.4
Child's Age (Months)	0.0	0.0	0.0	0.0	0.0	0	0.0	-	0.0	0.0	0.0	-	0	0.0

Characteristic	Sure Start: 7 years							Non Sure Start: 7 years						
	% missing by Strata (N=)					Missing for Strata 2-4 (N=)		% missing by Strata (N=)					Missing for Strata 2-4	
	Strata: SSLP Community Like							Strata: SSLP Community Like						
	Least				most	N	%	least				most	N	%
7 years	1	2	3	4	5			1	2	3	4	5		
Overall school effectiveness	5.3	22.3	15.1	15.1	13.4	580	16.3	-	7.1	15.5	6.7	-	140	9.7
Pupils attendance	5.3	22.5	15.9	15.2	13.4	594	16.7	-	7.5	15.5	7.6	-	146	10.2
% of pupils known to be eligible for free school meals	13.2	35.2	23.6	20.8	19.4	864	24.3	-	15.2	19.1	10.1	-	220	15.3
School CVA measure	7.9	39.0	31.8	29.4	20.9	1135	31.9	-	31.2	35.7	29.3	-	463	32.2

Table D.2: Percentage of data imputed- all seen at 7 years

Outcomes Measures	Outcome only							Outcome and demographic characteristics						
	% missing by Strata (N=7575)					For Strata 2-4 (N=4994)		% missing by Strata (N=4994)					For Strata 2-4	
	SS and MCS					Cases with outcomes	% missing	SS and MCS					Cases with outcomes	% missing
	SS 1	SS 2	SS 3	SS 4	SS 5			SS 1	2	3	4	SS 5		
BMI (standardised)	3.9	1.6	5.0	3.4	5.9	4819	3.5	-	13.2	18.6	18.7	-	2474	50.5
General health	1.3	0.1	0.4	0.3	0.2	4981	0.3	-	13.0	17.9	18.2	-	2543	49.1
Key stage 1 Reading	2.6	16.1	11.8	12.8	13.5	4332	13.3	-	13.9	18.3	18.7	-	2447	51.0
Key stage 1 Writing	2.6	16.1	11.8	12.8	13.5	4332	13.3	-	13.9	18.3	18.7	-	2447	51.0
Key stage 1 Maths	2.6	16.1	11.8	12.8	13.5	4332	13.3	-	13.9	18.3	18.7	-	2447	51.0
Key stage 1 Science	2.6	16.2	11.9	12.8	13.5	4328	13.3	-	13.9	18.3	18.7	-	2447	51.0
Emotional dysregulation	2.6	2.3	3.0	3.1	4.2	4850	2.9	-	13.2	18.2	18.7	-	2491	50.1
Positive social behaviour	2.6	1.6	2.1	2.0	2.1	4898	1.9	-	13.1	18.1	18.5	-	2511	49.7
Internalisation	2.6	2.0	2.8	2.5	3.9	4869	2.5	-	13.2	18.1	18.7	-	2497	50.0
Self regulation	1.3	1.4	1.4	1.9	1.8	4914	1.6	-	13.1	18.0	18.5	-	2516	49.6
Mother's depression	3.9	5.4	7.5	6.6	12.5	4664	6.6	-	3.6	5.4	5.3	-	4278	14.3

Outcomes Measures	Outcome only							Outcome and demographic characteristics						
	% missing by Strata (N=7575)					For Strata 2-4 (N=4994)		% missing by Strata (N=4994)					For Strata 2-4	
	SS and MCS					Cases with outcomes	% missing	SS and MCS					Cases with outcomes	% missing
	SS 1	SS 2	SS 3	SS 4	SS 5			SS 1	2	3	4	SS 5		
M's satisfaction with life	9.2	6.3	8.4	7.5	15.0	4617	7.5	-	3.8	5.8	5.6	-	4238	15.1
Harsh discipline in the home	2.6	3.1	3.6	2.8	4.2	4837	3.1	-	3.2	4.2	4.3	-	4410	11.7
Chaos in the home	1.3	0.1	0.4	0.5	0.5	4977	0.3	-	2.6	3.5	3.5	-	4515	9.6
Home learning environment	1.3	0.2	0.3	0.2	0.3	4982	0.2	-	2.6	3.4	3.5	-	4517	9.6

Appendix E: Producing measures of area characteristics

Area characteristics: a variety of census variables for each community (e.g. ethnic makeup, age distribution, employment status) and the 2004 IMD score were subjected to data-reduction-oriented factor analysis. Results were used to create composite factor scores reflecting dimensions of the community that could potentially influence the outcome measures. The labels of identified factors are listed in the left-hand column of Table E.1, with associated component variables defining each factor listed in the right-hand column.

Table E.1: Variables in area level composite factors

Composite	Variables in Composite
Economically deprived	High % lone parent families High % non working parents with children High % unemployed Low % all managerial Low % intermediate employment Low % small employers High % employment not classified Low % of all households owned High % all households social and council rented High % no qualifications High % people in households with no car or van High % household income < 60% national median High % of all people LLTI High IMD score 2004
Non Asian ethnic minority	High % Black African High % Black Caribbean High % Chinese High % mixed Low % white British High % white other Low % lower supervisory and technical Low % all routine employment Low % unshared of all occupied household spaces High % over 1.5 persons per room
Many children	Low % of all households with no dependent children High % of all people aged 0-4 Low % of all people aged 65+
Asian Pakistani	High % Asian Pakistani High % vacant household spaces
Transient population with children	High % inflow of all households with children High % outflow of all households with children
Asian Bangladeshi	High % Asian Bangladeshi
Asian Indian and students	High % Asian Indian High % economically active fulltime student

Appendix F: Effects of Strata and Covariates on Outcomes

In the results section the results of the multilevel models that tested for statistically significant effects upon outcomes are discussed primarily in terms of the main questions related to SSLP versus non-SSLP (MCS) differences. In analysing for such differences a large number of covariates (other predictor variables) were included in the multilevel models, so that SSLP effects could be determined after controlling for covariate differences. In the following tables of Appendix F, the effects of the covariates used in the multilevel models for each outcome are presented. A blank indicates no significant effect, -ve indicates a statistically significant ($p<.05$) negative effect (as covariate increases - outcome decreases) and +ve indicates a statistically significant ($p<.05$) positive effect (as covariate increases - outcome increases).

Results are presented firstly for analyses of the complete cases (F1), secondly for imputed data for all seen at 7 years (F2), and thirdly for imputed data for all seen at 3 years (F3).

Table F.1: Summary of Model Estimate Effects – 7 years: Complete data

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
SSLP (Baseline MCS)								+ve				+ve	-ve	-ve	+ve
Strata (Baseline Stratum 4)															
Stratum 2		-ve													
Stratum 3	-ve	-ve													
Child's Age					+ve										-ve
Childs Gender (Baseline Male)															
Female			+ve	+ve			-ve	+ve		+ve			-ve		+ve
Child's Ethnicity (Baseline White)															
Mixed															
Indian		-ve	-ve				-ve	+ve		+ve		+ve	-ve	-ve	
Pakistani		-ve			-ve										
Bangladeshi	+ve	-ve					-ve								
Black Caribbean															
Black Other		-ve			-ve						-ve	+ve			
Other							-ve	+ve							
Lang in Home (Baseline English)															
Other Languages							+ve								+ve
Maternal Age at Birth of Child (Baseline Not teenage)															
Teenage													+ve		+ve
Maternal Cognitive Difficulties (Baseline No Diffs)															
Some Difficulties				-ve		-ve								+ve	
Lone Parent (Baseline Not Lone)															
Lone Parent													+ve		

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Household Deprivation (Baseline Above poverty line[+)]															
Below poverty line													-ve		
Highest Education in Household (Baseline O level / GCSE)															
Degrees /Higher Education			+ve	+ve	+ve	+ve	-ve								
A level															
Other											+ve				
None		-ve			-ve			-ve					-ve		-ve
Highest Occupation in Household (Baseline Routine)															
Management/Prof			+ve	+ve	+ve	+ve		+ve							
Intermediate			+ve		+ve			+ve							
Small Employer					+ve										
Lower Super/Tech												+ve			
Semi-Routine															
Work Status Household (Baseline Working)															
Workless Household									+ve		+ve				
Area Variables															
Economically deprived				-ve		-ve									
Non Asian ethnic minority															
Many children		-ve	-ve		-ve		+ve	-ve	+ve		+ve		+ve		-ve
Asian Pakistani				+ve											
Transient population with children							+ve						+ve		
Asian Bangladeshi	+ve				+ve	+ve				-ve	-ve	+ve			-ve
Asian Indian and students	-ve	-ve				+ve	+ve	-ve				-ve	+ve	+ve	

	Outcome Variables														
	Physical Health		Child Educational Development				Child Behaviour and Social Development				Maternal Wellbeing		Parent and Family Functioning		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Urban/rural indicator (Baseline urban)															
Town															
Village		+ve		-ve			-ve								
Hamlet															
School variables															
Overall effectiveness						-ve							N/A	N/A	N/A
Pupil attendance			-ve	-ve	-ve	-ve	+ve	-ve	+ve				N/A	N/A	N/A
% free school meals													N/A	N/A	N/A
School CVA measure													N/A	N/A	N/A

Note: +ve = positive effect and –ve = negative effect

Unable to estimate

[+]Poverty line is £210 per week, taken from the financial year 2004-2005

Table F.2: Summary of Model Estimate Effects – all seen at 7 years: Imputed data

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
SSLP (Baseline MCS)								+ve	-ve			+ve	-ve	-ve	+ve
Strata (Baseline Stratum 4)															
Stratum 2								+ve							
Stratum 3															
Child's Age															-ve
Childs Gender (Baseline Male)															
Female			+ve	+ve			-ve	+ve	-ve	+ve			-ve		+ve
Child's Ethnicity (Baseline White)															
Mixed			+ve												
Indian	-ve	-ve					-ve			+ve			-ve	-ve	
Pakistani		-ve			-ve										
Bangladeshi															
Black Caribbean	+ve														
Black Other	+ve	-ve			-ve		-ve				-ve		-ve		
Other							-ve								
Lang in Home (Baseline English)															
Other Languages															+ve
Maternal Age at Birth of Child (Baseline Not teenage)															
Teenage													+ve		+ve
Maternal Cognitive Difficulties (Baseline No Diffs)															
Some Difficulties				-ve	-ve	-ve		-ve	+ve	-ve				+ve	
Lone Parent (Baseline Not Lone)															
Lone Parent													+ve		

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Household Deprivation (Baseline Above poverty line[+)]															
Below poverty line	-ve		-ve	-ve	-ve	-ve							-ve		
Highest Education in Household (Baseline O level / GCSE)															
Degrees /Higher Education	-ve		+ve	+ve	+ve	+ve	-ve							+ve	
A level													+ve		
Other	+ve						+ve			-ve	+ve				
None		-ve			-ve				+ve				-ve		-ve
Highest Occupation in Household (Baseline Routine)															
Management/Prof			+ve	+ve	+ve	+ve		+ve							
Intermediate			+ve		+ve			+ve							
Small Employer					+ve										
Lower Super/Tech															
Semi-Routine															
Work Status Household (Baseline Working)															
Workless Household									+ve	-ve	+ve				
Area Variables															
Economically deprived				-ve		-ve	+ve						+ve	+ve	
Non Asian ethnic minority	-ve														
Many children		-ve					+ve	-ve	+ve		+ve				-ve
Asian Pakistani	-ve					+ve									
Transient population with children													+ve		
Asian Bangladeshi	+ve									-ve		+ve			
Asian Indian and students	-ve						+ve							+ve	

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Urban/rural indicator (Baseline urban)															
Town		+ve		-ve	-ve										
Village			-ve	-ve		-ve	-ve								
Hamlet								-ve							
School variables															
Overall effectiveness	-ve		-ve	-ve	-ve	-ve							N/A	N/A	N/A
Pupil attendance			-ve		-ve			-ve	+ve				N/A	N/A	N/A
% free school meals													N/A	N/A	N/A
School CVA measure					+ve								N/A	N/A	N/A

Note: +ve = positive effect and –ve = negative effect

Unable to estimate

[+]Poverty line is £210 per week, taken from the financial year 2004-2005

Table F.3: Summary of Model Estimate Effects – all seen at 3 years: Imputed data

	Outcome Variables														
	Physical Health		Child Educational Development				Child Behaviour and Social Development				Maternal Wellbeing		Parent and Family Functioning		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
SSLP (Baseline MCS)											+ve		-ve	-ve	+ve
Strata (Baseline Stratum 4)															
Stratum 2													+ve		
Stratum 3															
Child's Age															-ve
Childs Gender (Baseline Male)															
Female			+ve	+ve			-ve	+ve	-ve	+ve			-ve		+ve
Child's Ethnicity (Baseline White)															
Mixed															
Indian		-ve					-ve						-ve		
Pakistani		-ve													
Bangladeshi															
Black Caribbean	+ve														
Black Other							-ve						-ve		
Other															
Lang in Home (Baseline English)															
Other Languages															+ve
Maternal Age at Birth of Child (Baseline Not teenage)															
Teenage							+ve						+ve		
Maternal Cognitive Difficulties (Baseline No Diffs)															
Some Difficulties			-ve	-ve	-ve	-ve	+ve	-ve	+ve	-ve				+ve	-ve
Lone Parent (Baseline Not Lone)															
Lone Parent													+ve		

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Household Deprivation (Baseline Above poverty line[+)]															
Below poverty line	-ve		-ve	-ve	-ve	-ve							-ve	+ve	
Highest Education in Household (Baseline O level / GCSE)															
Degrees /Higher Education			+ve	+ve	+ve	+ve	-ve		-ve						+ve
A level															
Other	+ve										+ve				
None		-ve	-ve	-ve	-ve	-ve			+ve				-ve		-ve
Highest Occupation in Household (Baseline Routine)															
Management/Prof			+ve	+ve	+ve	+ve									
Intermediate			+ve	+ve	+ve										
Small Employer															
Lower Super/Tech															
Semi-Routine															
Work Status Household (Baseline Working)															
Workless Household								-ve	+ve		+ve				
Area Variables															
Economically deprived			-ve	-ve	-ve	-ve	+ve						+ve	+ve	
Non Asian ethnic minority															
Many children															-ve
Asian Pakistani															
Transient population with children				-ve											
Asian Bangladeshi										-ve	-ve	+ve		-ve	
Asian Indian and students							-ve							+ve	

	Outcome Variables														
	Physical Health		**Child Educational Development**				**Child Behaviour and Social Development**				**Maternal Wellbeing**		**Parent and Family Functioning**		
Demographic Variables	BMI (standardised)	General health	Key stage 1 Reading Attainment	Key stage 1 Writing Attainment	Key stage 1 Maths Attainment	Key stage 1 Science Attainment	Emotional dysregulation	Positive social behaviour	Internalisation	Self regulation	Mother's depression	M's satisfaction with life	Harsh discipline in the home	Chaos in the home	Home learning environment
Urban/rural indicator (Baseline urban)															
Town															
Village			-ve	-ve											
Hamlet															
Ofsted school variables															
Overall effectiveness			-ve	-ve	-ve	-ve							N/A	N/A	N/A
Pupil attendance			-ve		-ve			-ve					N/A	N/A	N/A
% free school meals							+ve						N/A	N/A	N/A
Pupil CVA measure													N/A	N/A	N/A

Note: +ve = positive effect and –ve = negative effect

Unable to estimate

[+]Poverty line is £210 per week, taken from the financial year 2004-2005

Appendix G: SSLP vs. MCS by demographic group interactions

Note that estimated values are derived from models using interactions of interest (baseline characteristics assumed for all other variables).

In section 3.3 the results of analyses that considered whether there were differences in SSLP effects for different demographic subgroups were discussed. In Appendix G tables are presented that show the results where there was a significant interaction between SSLP status and subgroup status. Where the p-value is less than .05 then this indicates a statistically significant interaction between SSLP/MCS status and subgroup status.

The results are presented firstly for analyses of the imputed data for all seen at 7 years.

Appendix G: SSLP status by demographic group interactions – significant results

Note that estimated values are derived from models using interactions of interest (baseline characteristics assumed for all other variables).

Table G1: SSLP status by demographic interactions - significant results with p values: Imputed 7 yr data

Outcome Measure	Categories	Sure Start		Non Sure Start		P-value
		Estimated values	95% CI	Estimated values	95% CI	
M's Satisfaction with life	**Lone Parent**					
	Not Lone Parent	0.17	-0.86 to 1.21	0.16	-0.87 to 1.20	0.79
	Lone Parent	0.18	-0.87 to 1.24	-0.06	-1.14 to 1.02	0.0002
	Work status in HH					
	Working	0.14	-0.88 to 1.16	0.16	-0.87 to 1.18	0.66
	Workless	0.14	-0.88 to 1.14	-0.14	-1.17 to 0.90	0.0006
Chaos	**Child's Gender**					
	Male	-0.58	-1.57 to 0.42	-0.17	-1.17 to 0.84	<0.0001
	Female	-0.53	-1.53 to 0.46	-0.36	-1.44 to 0.71	0.09

Appendix G: SSLP vs. MCS by demographic group interactions – significant results

Note that estimated values are derived from models using interactions of interest (baseline characteristics assumed for all other variables).

Table G.2: SSLP status by demographic interactions - significant results with p values: Progress Imputed 7 yr data

Outcome Measure	Categories	Sure Start		MCS		P-value
		Estimated values		Estimated values		
		In sd units Effect size	(95% CI)	In sd units Effect size	(95% CI)	
M's Satisfaction with life	**Lone Parent**					
	Not Lone Parent	-0.68	-1.79 to 0.42	-0.69	-1.82 to 0.43	0.75
	Lone Parent	-0.64	-1.77 to 0.48	-0.84	-2.02 to 0.33	0.007
	Work status in HH					
	Working	-0.71	-1.79 to 0.38	-0.70	-1.80 to 0.41	0.83
	Workless	-0.72	-1.77 to 0.33	-0.95	-2.05 to 0.15	0.003

List of reports from the National Evaluation of Sure Start, 2001-2012

Available at www.ness.bbk.ac.uk

Implementation Reports

Details of the Implementation Evaluation Proposal .

September 2001 Implementation Methodology Report

Full report of Buildings in Sure Start Local Programmes

Full Report on Early Experiences of Implementing Sure Start.

Full Report of the Empowering Parents Report

Full Report on Fathers in Sure Start Local Programmes

Full Report on Getting Sure Start Started.

Full report on Improving the employability of parents in Sure Start local programmes

Full report Implementing Sure Start local programmes: An in-depth study.

Full report Implementing Sure Start local programmes: An in-depth study. Part two - A close up on services.

Full report of Implementing Sure Start Local Programmes: An Integrated Overview.

Full report of Maternity Services in Sure Start Local Programmes

Full Report of the Outreach and Home Visiting Report

Full report of The Quality of Early Learning, Play and Childcare Services in Sure Start Local Programmes

Full report: Sure Start Local Programmes work on Domestic Abuse

Full report: A Better Start: Children and Families with Special Needs and Disabilities in Sure Start Local Programmes

Full report: Sure Start and Black and Minority Ethnic Populations

Full report: Promoting Speech and Language – a themed study in fifteen Sure Start Local Programmes

Full report: Family and Parenting Support in Sure Start Local Programmes

Full report: Understanding the contribution of Sure Start Local Programmes to the task of safeguarding children's welfare

Impact Study reports

Details of the Implementation Evaluation Proposal .

Impact Study Methodology Report September 2002

Early Impacts of SSLPs on Children & Families for the Full Report. - 2005

Variation in SSLP Effectiveness: Early Preliminary Findings for the Full Report. - 2005

Understanding Variations in Effectiveness amongst Sure Start Local Programmes. -2007

The Impact of Sure Start Local Programmes on Three Year Olds and Their Families. -2008

The Impact of Sure Start Local Programmes on Five Year Olds and Their Families. -2010

The quality of group childcare settings used by 3-4 year old children in Sure Start Local Programme areas and the relationship with child outcomes. -2010

Report of the Longitudinal Study of 7-year-old Children and their Families -2012

Local Context Analysis

Local Context Analysis Evaluation

Local Context Analysis Methodology Report

Full Report on Characteristics of Sure Start local programme areas: Rounds 1 to 4

Full report on Characteristics of Sure Start local programmes - 2001/2

Full Report of Changes in the Characteristics of SSLP Areas - 2000/2001 to 2002/2003

Full Report of Changes in the Characteristics of SSLP Areas - 2000/1 to 2003/4

Full Report: Changes in the Characteristics of SSLP Areas - 2000/1 to 2004/5

Cost Effectiveness Evaluation

Cost Effectiveness Evaluation

Cost Effectiveness Methodology Report

Cost Effectiveness of Implementing SSLPs: An Interim Report

Support to Local Programmes

Methodological Progress Report

Programme Evaluations

Cost Effectiveness in Sure Start Local Programmes: A Synthesis of Local Evaluation Findings

Smoking Cessation Services in SS Local Programmes: Findings from Local Evaluations

Black & Minority Ethnic Families & Sure Start: Findings from Local Evaluation Reports

Breastfeeding , Weaning and Health Eating

Full report on Partnership Working in Sure Start Local Programmes

Speech and Language Synthesis Report

Implementing and Managing your Sure Start Local Programme Evaluation

Conducting Ethical Research.

Guidance on Local Cost-Effectiveness.

Guidance on Local Evaluation - Sure Start Unit.

Conducting User Satisfaction Surveys.

Involving Parents and Carers Guidance document

Not by NESS but on website as useful Exploring the Field of Listening to and Consulting with Young Children document

Using Existing Data in Sure Start Local programme evaluations

Sharing Evaluation Findings: Disseminating the Evidence

Qualitative Research Methods

Measuring Outcomes: Guidance on Outcome Evaluation for SSLPs

Ref: DFE-RR220

ISBN: 978-1-78105-116-0

June 2012